MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG VON

E. KAMKE
TÜBINGEN

K. KNOPP
TÜBINGEN

R. NEVANLINNA
HELSINKI

E. SCHMIDT
BERLIN

F. K. SCHMIDT
HEIDELBERG

HERAUSGEGEBEN VON

H. WIELANDT
TÜBINGEN

WISSENSCHAFTLICHER BEIRAT

W. BLASCHKE L. FEJÉR A. E. INGHAM
H. KNESER W. MAGNUS O. PERRON
W. SÜSS H. WEYL

63. BAND, 1. HEFT

(ABGESCHLOSSEN AM 6. SEPTEMBER 1955)

BERLIN · GÖTTINGEN · HEIDELBERG
SPRINGER-VERLAG
1955

Math. Z.

Die **MATHEMATISCHE ZEITSCHRIFT**

wurde im Jahre 1918 von *L. Lichtenstein* unter der Mitwirkung von *K. Knopp*, *E. Schmidt* und *I. Schur* gegründet und herausgegeben. Nach dem Tode Lichtensteins übernahm *K. Knopp* 1933 die Herausgabe. Die Schriftleitung ergänzte sich 1933 durch *E. Kamke* und *F. K. Schmidt*, 1936 durch *R. Nevanlinna* und 1950 durch *H. Wielandt*, der 1952 die Herausgabe übernahm.

Die Zeitschrift erscheint nach Maßgabe des eingehenden Materials. Der Preis des Bandes beträgt DM 96.—.

Die Mathematische Zeitschrift ist durch jede Buchhandlung zu beziehen.

Die Mathematische Zeitschrift dient der Pflege der reinen Mathematik, doch werden auch Beiträge aus den Gebieten der theoretischen Physik und Astronomie Aufnahme finden, soweit sie mathematisch von Interesse sind. Besprechungen, Aufgaben u. dgl. werden nicht zugelassen.

Es wird ausdrücklich darauf aufmerksam gemacht, daß mit der Annahme des Manuskriptes und seiner Veröffentlichung durch den Verlag das ausschließliche Verlagsrecht für alle Sprachen und Länder an den Verlag übergeht, und zwar bis zum 31. Dezember desjenigen Kalenderjahres, das auf das Jahr des Erscheinens folgt. Hieraus ergibt sich, daß grundsätzlich nur Arbeiten angenommen werden können, die vorher weder im Inland noch im Ausland veröffentlicht worden sind, und die auch nachträglich nicht anderweitig zu veröffentlichen der Autor sich verpflichtet.

Die Mitarbeiter erhalten von ihren Arbeiten 75 Sonderdrucke unentgeltlich.

Manuskriptsendungen sind zu richten an die Mitglieder der unterzeichneten Redaktion. Die Herren Mitarbeiter werden im Interesse einer raschen Drucklegung gebeten, die Arbeiten in gut lesbarer Niederschrift einzureichen. Für den Text ist, wenn irgend möglich, Maschinenschrift, für die mathematischen Formeln jedoch nur Handschrift zu verwenden. Etwaige Abbildungen sind auf einem besonderen Blatt zu zeichnen. Die in den Formeln etwa vorkommenden griechischen oder deutschen Buchstaben (Fraktur) sowie andere der Gefahr von Verwechselungen ausgesetzte Zeichen, z. B. die Landauschen Symbole o und O, sind stets besonders (etwa mit verschiedenfarbigen Stiften) zu kennzeichnen. Die Fußnoten sind fortlaufend zu numerieren, bei Zitaten Erscheinungsjahr und Seitenzahlen anzugeben

Wenn zu einer bereits gedruckten Arbeit eine Berichtigung nötig wird, wollen die Herren Verfasser dem Herausgeber hiervon unverzüglich Mitteilung machen. Die Berichtigungen sollen dann am Schlusse eines Bandes gesammelt aufgenommen werden.

REDAKTION DER MATHEMATISCHEN ZEITSCHRIFT:

E. Kamke, Tübingen, Frischlinstraße 27,
K. Knopp, Tübingen, Am Apfelberg 6,
R. Nevanlinna, Helsinki (Finnland), Mäntytie 5 B,
E. Schmidt, Berlin-Steglitz, Sedanstraße 8,
F. K. Schmidt, Heidelberg, Mathematisches Institut
H. Wielandt, Tübingen, Biesingerstraße 7.

Springer-Verlag OHG — Berlin W 35, Reichpietschufer 20. — Heidelberg, Neuenheimer Landstraße 24. — Göttingen, Weender Straße 60.

---

63. Band — **Inhalt** — 1. Heft

Seite

Issai Schur zum Gedächtnis 1
Ostrowski, A., Über Normen von Matrizen 2
Perron, O., Über die Abhängigkeit von Potenzsummen und einen Satz von Pólya . 19
Rosenthal, A., On the continuity of functions of several variables 31
Knopp, K., Nörlund-Verfahren für Funktionen. (Zweite Mitteilung) 39
Dieudonné, J., Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$. III 53
Neumann, B. H., Groups with finite classes of conjugate subgroups 76
Rogosinski, W. W., On finite systems of linear equations with an infinity of unknowns 97
Dinghas, A., Konvexitätseigenschaften von Mittelwerten harmonischer und ververwandter Funktionen 109

Issai Schur zum Gedächtnis.

Aus Anlaß der 80. Wiederkehr des Tages, an dem SCHUR in Mohilew am Dnjepr geboren wurde, vereinen sich Freunde und Schüler, um sein Andenken mit diesem Bande der Zeitschrift zu ehren, die er selbst begründet hat. Sie sind in alle Welt zerstreut durch die Katastrophe, in deren Verlauf SCHUR durch vorzeitige Emeritierung 1935 die Wirkungsstätte verlor, an der seine Vorlesungen drei Jahrzehnte lang Studenten für die Mathematik begeistert hatten. SCHUR konnte den furchtbarsten Folgen entkommen, indem er 1939 von Berlin nach Israel übersiedelte. Doch war es ihm nicht vergönnt, die wissenschaftliche Tätigkeit wieder aufzunehmen. Als er in Israel eintraf, war er bereits von dem Herzleiden gezeichnet, dem er am 10. Januar 1941, seinem 66. Geburtstag, in Tel Aviv erlag.

SCHUR hat, wie seine Antrittsrede in der Preußischen Akademie der Wissenschaften (1922) erkennen läßt, seine Aufgabe darin gesehen, die Algebra insbesondere auf dem Gebiet der linearen Transformationen zu fördern und in eine lebendige Wechselwirkung mit der Analysis zu bringen. Die Erfüllung dieser Aufgabe hat ihn als Mitbegründer der Darstellungstheorie in die Geschichte der Mathematik eingehen lassen. Doch enthalten auch seine zahlreichen auf andere Fragen gerichteten, stets zu vollendeter Form geschliffenen Arbeiten eine Fülle von Gedanken, die noch immer nicht ausgeschöpft ist. Möge dieser Band an SCHURs Gesamtwerk erinnern und von seiner Fruchtbarkeit zeugen.

Die Schriftleitung.

Ostrowski, A.
Math. Zeitschr. Bd. 63, S. 2—18 (1955)

Über Normen von Matrizen.

Dem Andenken an Issai Schur gewidmet.

Von

Alexander Ostrowski.

Einleitung.

1. Im folgenden sollen die großen lateinischen Buchstaben $A, B, C, \ldots$ mit Ausnahme des § 3 quadratische Matrizen einer festen Ordnung n mit reellen oder komplexen Elementen $a_{\mu\nu}$ bzw. $b_{\mu\nu}, c_{\mu\nu}, \ldots$ bezeichnen. Werden andere Matrizen betrachtet, so wird dies jedesmal ausdrücklich angegeben werden. Die griechischen Buchstaben $\xi, \eta, \ldots$ bedeuten n-dimensionale (Kolonnen-) Vektoren mit den Komponenten $x_1, x_2, \ldots, x_n$ bzw. $y_1, y_2, \ldots, y_n, \ldots$.

2. Bei verschiedenartigsten Untersuchungen wurden bereits zur ungefähren Charakterisierung der „Größe" einer Matrix verschiedene Ausdrücke benutzt, die man als *Beträge, Schranken* oder auch *Normen* bezeichnet. In der Literatur sind vor allem sechs verschiedene solche Normen verwendet worden. Es sind dies in der weiter unten zu erläuternden Nomenklatur die drei *Hölderschen Normen*

$$(1) \qquad |A|_1 = \sum_{\mu,\nu} |a_{\mu\nu}|, \qquad |A|_2 = \left(\sum_{\mu,\nu} |a_{\mu\nu}|^2\right)^{\frac{1}{2}}, \qquad N^*(A) = |A|_\infty = \max_{\mu,\nu} |a_{\mu\nu}|,$$

sowie die zu den drei entsprechend definierten Vektornormen gehörenden symmetrischen *Schrankennormen*.

Es scheint nunmehr angebracht zu sein, eine systematische Untersuchung des Normbegriffes vorzunehmen unter Zugrundelegung geeigneter Postulate[1]). Sind auch, wie es sich herausstellt, für Limes- und Konvergenzbetrachtungen alle hier in Frage kommenden Normbegriffe im wesentlichen (für festes n) gleichberechtigt, so liefern sie andererseits in der Abschätzung der Eigenwerte oder bei der Diskussion des Fehlers in verschiedenen Iterationsverfahren sehr verschiedene Ergebnisse, so daß es wertvoll erscheint, über eine genügende Anzahl Normen zu verfügen. Dazu kommt noch, daß nicht alle Normen für Grenzübergänge mit $n \to \infty$ brauchbar sind. Es scheint, daß hierzu nur die symmetrischen Schrankennormen verwendet werden können.

[1]) Diese Untersuchung wurde für die sog. Schrankennormen in einer Note des Verfassers (Ostrowski [2]) begonnen und in allgemeineren Fällen von Herrn We. Gautschi in seiner Basler Dissertation ([2], [3]) fortgesetzt.

3. Wir definieren im folgenden eine *allgemeine Norm* $N(A)$ durch die drei Postulate:

I_N $N(A) > 0$ $(A \neq 0)$;

II_N $N(cA) = |c| N(A)$ für jede skalare Größe c;

III_N $N(A + B) \leq N(A) + N(B)$ [2]).

Aus II_N folgt offenbar für $c = 0$, daß $N(0) = 0$ ist.

4. Neben diesen Normen betrachten wir noch speziellere, sog. *multiplikative Normen* $M(A)$, die den folgenden vier Postulaten genügen (vgl. z.B. FADDEJEWA [1], wo aber I_N statt I_M benutzt wird, p. 60; NBSR p. 83):

I_M $M(A) \geq 0$ und für wenigstens ein A_0: $M(A_0) > 0$;

II_M $M(cA) = |c| M(A)$ für jede skalare Größe c;

III_M $M(A + B) \leq M(A) + M(B)$;

IV_M $M(AB) \leq M(A) M(B)$.

5. Wie man sieht, kommt hier das Postulat IV_M dazu; andererseits wird das Postulat I_M abgeschwächt. Wir beweisen nun im § 1, der der Diskussion der multiplikativen Normen gewidmet ist, den

Satz I. *Jede multiplikative Norm ist auch eine allgemeine Norm.*

Andererseits ergibt sich, daß $M(E)$, unter E die Einheitsmatrix verstanden, stets ≥ 1 ist. Offenbar ist keine ähnliche Relation bei allgemeinen Normen generell möglich, da ja mit $N(A)$ auch $cN(A)$ für jede positive Konstante c eine allgemeine Norm ist.

6. Wir bezeichnen in dieser Arbeit mit λ_A den maximalen absoluten Betrag der Fundamentalwurzeln der Matrix A. Eine besonders wichtige Eigenschaft multiplikativer Normen kommt zum Ausdruck im

Satz II. *Es gilt stets* $M(A) \geq \lambda_A$ [3]).

7. Für einen allgemeinen Vektor ξ wird eine allgemeine Norm $h(\xi)$ durch die Postulate definiert (vgl. z.B. FADDEJEWA [1], p. 59; NBSR p. 81):

I_V $h(\xi) > 0$ $(\xi \neq 0)$;

II_V $h(c\xi) = |c| h(\xi)$ für jede skalare Größe c;

III_V $h(\xi + \eta) \leq h(\xi) + h(\eta)$.

8. Sind nun $h(\xi)$ und $h_1(\xi)$ zwei derartige Vektornormen, so wird durch

$$S_{h,h_1}(A) = \operatorname{Sup}_{\xi \neq 0} \frac{h(A\xi)}{h_1(\xi)} \tag{2}$$

eine *Schrankennorm* von A definiert, die zu den Vektornormen $h(\xi)$ und $h_1(\xi)$ gehört. Ist insbesondere $h(\xi) = h_1(\xi)$, so haben wir es mit einer

[2]) Diese Postulate sind von GAUTSCHI [2], p. 375 als Postulate VI, VII und VIII benutzt worden.

[3]) Für zahlreiche Spezialfälle und auch für einige nicht multiplikative Normen hat GAUTSCHI ([1], [2], [3]) die Ungleichung des Satzes II bewiesen. Im allgemeinsten Falle einer multiplikativen Norm ist der Satz II auf einem anderen Wege als bei uns bewiesen bei FADDEJEWA [1], p. 65; NBSR p. 90.

symmetrischen Schrankennorm zu tun (vgl. z.B. FADDEJEWA [*1*], pp. 60, 61; NBSR pp. 84, 85):

$$S_h(A) = \operatorname*{Sup}_{\xi \neq 0} \frac{h(A\xi)}{h(\xi)}. \tag{3}$$

Offenbar gilt für symmetrische Schrankennormen stets

$$S_h(E) = 1. \tag{4}$$

Übrigens folgt aus den Postulaten I_V bis III_V, daß eine Vektornorm stets eine *stetige Funktion der Vektorkomponenten* ist (vgl. den Stetigkeitsbeweis für allgemeine Normen in Nr. 20). Daher kann man in den Definitionen (2) und (3) das *Supremum* durch das *Maximum* ersetzen.

9. Aus der Definition (2) und den Postulaten I_V bis III_V folgt sofort, daß $S_{h,h_1}(A)$ den Postulaten I_N bis III_N genügt, also eine allgemeine Norm ist. Für die symmetrische Schrankennorm $S_h(A)$ folgt aber leicht, daß auch das Postulat IV_M erfüllt ist, so daß $S_h(A)$ stets eine *multiplikative Norm* ist. In der Tat gilt ja

$$\frac{h(AB\xi)}{h(\xi)} = \left(\frac{h(AB\xi)}{h(B\xi)}\right)\left(\frac{h(B\xi)}{h(\xi)}\right),$$

woraus IV_M sofort folgt. Im allgemeinsten Falle der nicht notwendig symmetrischen Schrankennormen wird in Nr. 19 als Formel (1.10) eine anders beschaffene „Multiplikationseigenschaft" aufgestellt.

10. Ist $N(A)$ eine allgemeine Matrizennorm, so läßt sich aus ihr eine Vektornorm $h(\xi)$ herleiten, indem man dem Vektor $\xi = (x_1, \ldots, x_n)$ eine $(n \times n)$-Matrix A_ξ zuordnet, die in der ersten Kolonne $x_1, x_2, \ldots, x_n$ hat und in den übrigen Kolonnen lauter Nullen. Dann kann man definieren

$$h_N(\xi) = N(A_\xi)^{4}). \tag{5}$$

Am Schlusse des § 1 beweisen wir den

Satz III. *Wird aus einer multiplikativen Norm $M(A)$ die zugehörige Vektornorm $h_M(\xi)$ hergeleitet und sodann die zu dieser Vektornorm gehörende symmetrische Schrankennorm $S_{h_M}(A)$ gebildet, so gilt stets*

$$M(A) \geq S_{h_M}(A). \tag{6}$$

11. Im § 2, der allgemeinen Normen $N(A)$ gewidmet ist, beweisen wir zuerst den

Satz IV. *Für zwei beliebige Normen $N(A)$ und $N_1(A)$ liegt der Quotient $\frac{N(A)}{N_1(A)}$ zwischen zwei positiven von A unabhängigen Konstanten.*

Es genügt hierzu zu beweisen, daß

$$p_1 < \frac{N(A)}{N^*(A)} < p_2 \tag{7}$$

[4]) Natürlich könnte man hier auch $N(A_\xi)$ so definieren, daß die Komponenten von A_ξ anstatt in der *ersten* Kolonne in der *k-ten* Kolonne von A_ξ stehen, und man erhält auf diese Weise mehrere aus $N(A)$ herleitbare Vektornormen, auf die unsere Betrachtungen genau so anwendbar bleiben.

gilt, wo $N^*(A)$ die in (1) definierte Norm ist und p_1 und p_2 von A unabhängige positive Zahlen sind.

Sodann beweisen wir eine Abschätzung von $N(A^t)$ für natürliche t, die für $t \to \infty$ die „richtige" Größenordnung von $N(A^t)$ liefert. Wir leiten nämlich her den

Satz V. *Es seien die Elementarteiler der Matrix A durch*

$$(\lambda - \lambda_1)^{m_1},\ (\lambda - \lambda_2)^{m_2},\ \ldots,\ (\lambda - \lambda_k)^{m_k} \tag{8}$$

gegeben, wobei die Anordnung so ist, daß

$$\lambda_A = |\lambda_1| = \cdots = |\lambda_r| > |\lambda_{r+1}| \geq \cdots \geq |\lambda_k| \tag{9}$$

und ferner

$$M \equiv m_1 = \cdots = m_s > m_{s+1} \geq \cdots \geq m_r \tag{10}$$

gilt. Dann besteht für jede allgemeine Norm $N(A)$ die Relation

$$p_3 < \frac{N(A^t)}{t^{M-1}\lambda_A^t} < p_4, \tag{11}$$

wo p_3 und p_4 zwei positive von t unabhängige und nur von $N(A)$ abhängende Zahlen sind. Gilt aber

$$\lambda_1 = \lambda_2 = \cdots = \lambda_s, \tag{12}$$

so konvergiert der Quotient in (11) sogar gegen einen positiven endlichen Grenzwert[5]).

Aus den Sätzen IV und V folgt dann sofort der

Satz VI. *Für jede allgemeine Norm gilt*

$$\sqrt[t]{N(A^t)} \to \lambda_A \qquad (t \to \infty)\,[6]). \tag{13}$$

12. Im § 3 diskutieren wir die *Hölderschen Matrizennormen*

$$|A|_p = \Big(\sum_{\mu,\nu} |a_{\mu\nu}|^p\Big)^{1/p} \qquad (p \geq 1). \tag{14}$$

---

[5]) Die aus (11) folgende Abschätzung der allgemeinen Norm $N(A)$ *nach oben* hat in etwas schwächerer Form Herr Gautschi ([*1*], [*2*]) bewiesen [dort ist M das Maximum sämtlicher Exponenten m_k in (8)]. Die zweite, unter der Annahme (12) geltende Behauptung des Satzes V hatte ich ursprünglich nur für die spezielle Norm $N^*(A)$ bewiesen. Erst durch Herrn Gautschi bin ich in freundlicher Weise darauf aufmerksam gemacht worden, daß eine geringfügige Abänderung meines Beweises sie für jede allgemeine Norm liefert.

[6]) Die Relationen (7) und (13) sind unabhängig von mir auch von Herrn L. Carleson bewiesen worden im Laufe einer für mich sehr anregenden Diskussion, für die ich ihm auch hier meinen herzlichsten Dank aussprechen möchte.

Es sei noch bemerkt, daß die Relation (13) für den allgemeinsten Fall von (symmetrischen und unsymmetrischen) Schrankennormen sich bei Gautschi [*2*], p. 377, Formel (34), findet. Im Falle der symmetrischen Schrankennorm stellt sie einen Spezialfall eines allgemeinen Satzes der Theorie der linearen Transformationen im Hilbertschen Raum dar (vgl. Riesz-Nagy [*1*], p. 421).

Bei dieser Diskussion werden wir die Annahme, daß die Matrizen $A, B, C, \ldots$ quadratische Matrizen sind, fallen lassen und sie für diesen ganzen Paragraphen nur als rechteckige Matrizen voraussetzen, wobei natürlich, damit das Produkt AB „existiert", vorausgesetzt werden muß, daß die Kolonnenanzahl von A gleich der Zeilenzahl von B ist. Wir beweisen dann den Satz VII, aus dem insbesondere folgen wird, daß *im Falle quadratischer Matrizen die Norm* $|A|_p$ *für* $1 \leq p \leq 2$ *eine multiplikative Norm ist*:

Satz VII. *Existiert das Matrizenprodukt* AB *und ist* $1 \leq p \leq 2$, *so gilt*[7])

$$|AB|_p \leq |A|_p |B|_p \qquad (1 \leq p \leq 2), \tag{15}$$

während (15) *für jedes* $p > 2$ *dann und nur dann allgemein*[8]) *gilt, wenn* A *aus einer einzigen Zeile und* B *aus einer einzigen Kolonne besteht.*

Für $p \geq 2$ gelten aber zwei der Hölderschen Ungleichung analoge Relationen.

Satz VIII. *Ist* $p \geq 2$, $\frac{1}{p} + \frac{1}{q} = 1$ *und existiert das Produkt* $C = AB$, *so gilt*

$$|AB|_p \leq |A|_p |B|_q, \tag{16}$$

$$|AB|_p \leq |A|_q |B|_p. \tag{17}$$

Dagegen gilt für jedes p *mit* $1 \leq p < 2$ *die Relation* (16) *dann und nur dann allgemein, wenn die Matrix* B *aus einer einzigen Kolonne besteht. Analog gilt* (17) *für jedes* p *mit* $1 \leq p < 2$ *dann und nur dann allgemein, wenn die Matrix* A *aus einer einzigen Zeile besteht.*

13. Im Falle einer *ungeraden* Anzahl von Faktoren gilt aber eine ähnliche Ungleichung für *alle* $p \geq 1$:

Satz IX. *Ist* $u > 1$ *ungerade, und existiert das Produkt* $A_1 A_2 \ldots A_u$, *so gilt für* $\frac{1}{p} + \frac{1}{q} = 1$

$$|A_1 A_2 \ldots A_u|_p \leq |A_1|_p |A_2|_q |A_3|_p \ldots |A_u|_p \qquad (p \geq 1). \tag{18}$$

Ist aber $g > 0$ *gerade und sind die Matrizen* $A_1, \ldots, A_g$ *die allgemeinsten Matrizen mit den vorgegebenen Zeilen- und Kolonnenanzahlen, für die das Produkt* $A_1 \ldots A_g$ *existiert, so gilt allgemein die Relation*

$$|A_1 A_2 \ldots A_g|_p \leq |A_1|_p |A_2|_q \ldots |A_{g-1}|_p |A_g|_q \tag{19}$$

dann und nur dann, wenn entweder $p \geq 2$ *oder die Kolonnenanzahl in* A_g *gleich* 1 *ist.*

Über die zu (17) analoge Ungleichung gilt der

Satz X. *Sind* $A_1, \ldots, A_l$ l *allgemeinste Matrizen mit vorgegebenen Zeilen- und Kolonnenanzahlen, deren Produkt existiert, so gilt im Falle eines ungeraden* l

$$|A_1 A_2 \ldots A_l|_p \leq |A_1|_q |A_2|_p \ldots |A_l|_q \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right) \tag{20}$$

[7]) Für $p = 2$ wurden diese Ungleichungen schon seit längerer Zeit verwendet (vgl. OSTROWSKI [1]).

[8]) Wir sagen hier (und später), daß eine solche Relation „allgemein" gilt, wenn sie für *alle* Matrizen mit festen Zeilen- und Kolonnenanzahlen gilt.

dann und nur dann allgemein, wenn entweder $p \geq 2$ oder die Zeilenanzahl in A_1 und die Kolonnenanzahl in A_l beide gleich 1 sind. Für ein gerades g gilt aber

$$|A_1 A_2 \dots A_g|_p \leq |A_1|_q |A_2|_p \dots |A_g|_p \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right) \tag{20°}$$

allgemein dann und nur dann, wenn entweder $p \geq 2$ oder die Zeilenanzahl in A_1 gleich 1 ist.

— Übrigens ergeben sich im Text die Beweise der Unmöglichkeitsbehauptungen der Sätze VII bis X, indem man sämtliche Elemente der Matrizen gleich 1 setzt.

Unsere Beweise der obigen Sätze machen fortgesetzt Gebrauch von der folgenden speziellen Norm der Matrix A:

$$|A|_{q,p} = \left[\sum_\nu \left(\sum_\mu |a_{\mu\nu}|^q\right)^{p/q}\right]^{1/p} \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right), \tag{21}$$

über die im § 3 verschiedene Hilfsformeln (3.9), (3.10), (3.11) und (3.12) hergeleitet werden. Diese spezielle Norm kann aber auch bei gewissen Abschätzungen von Matrizenpotenzreihen verwendet werden. Wir beweisen daher noch den

Satz XI. *Ist A eine quadratische Matrix, so gilt für jedes natürliche l*

$$|A^l|_p \leq |A|_p |A|_{q,p}^{l-1} \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right). \tag{22}$$

Sind allgemeiner $A_1, \dots, A_l$ rechteckige Matrizen, deren Produkt $A_1 \dots A_l$ existiert, so gilt

$$|A_1 \dots A_l|_p \leq |A_1'|_{q,p} \dots |A_{k-1}'|_{q,p} |A_k|_p |A_{k+1}|_{q,p} \dots |A_l|_{q,p}, \tag{23}$$

wo k auch gleich 1 oder l sein kann.

Endlich geben wir zum Schluß des § 3 verschiedene vereinzelte Formeln für die Schrankennormen, die zu den Hölderschen Vektornormen gehören.

§ 1. Multiplikative Normen.

14. Aus I_M und IV_M folgt für $A = A_0$ und $B = E$, wo E die Einheitsmatrix bedeutet, $M(A_0) \leq M(A_0) M(E)$,

$$M(E) \geq 1. \tag{1.1}$$

15. Wir verstehen im folgenden unter $E_{\mu\nu}$ die Matrix, bei der das in der μ-ten Zeile und der ν-ten Kolonne stehende Element gleich 1 ist, während alle übrigen Elemente verschwinden. Dann gelten für jedes $A = (a_{\mu\nu})$ die allgemeinen Formeln

$$A = \sum_{\sigma,\varkappa} a_{\sigma\varkappa} E_{\sigma\varkappa}, \tag{1.2}$$

$$E_{\mu\nu} A = \sum_\varkappa a_{\nu\varkappa} E_{\mu\varkappa}. \tag{1.3}$$

Um (1.3) aus (1.2) herzuleiten, genügt es zu beachten, daß allgemein, unter $\delta_{\sigma\nu}$ das Kroneckersche Symbol verstanden, gilt

$$E_{\mu\nu} E_{\sigma\varkappa} = \delta_{\nu\sigma} E_{\mu\varkappa}. \tag{1.4}$$

Multiplizieren wir (1.3) rechtsseitig mit $E_{\sigma\mu}$ und beachten (1.4) für $\varkappa = \mu$, so folgt, wegen $E_{\mu\varkappa}E_{\sigma\mu} = \delta_{\varkappa\sigma}E_{\mu\mu}$,

$$E_{\mu\nu}AE_{\sigma\mu} = a_{\nu\sigma}E_{\mu\mu}. \tag{1.5}$$

Ersetzen wir in (1.5) A durch $E_{\nu\nu}$ und setzen $\sigma = \nu$, so ergibt sich

$$E_{\mu\nu}E_{\nu\nu}E_{\nu\mu} = E_{\mu\mu}. \tag{1.6}$$

16. Wegen $\sum_{\nu=1}^{n} E_{\nu\nu} = E$, wo E die Einheitsmatrix ist, folgt nach III_M und (1.1)

$$\sum_{\nu=1}^{n} M(E_{\nu\nu}) > 0.$$

Daher ist $M(E_{\mu\mu})$ für wenigstens ein μ positiv. Aus (1.6) folgt aber wegen IV_M

$$M(E_{\mu\mu}) \leqq M(E_{\mu\nu})\, M(E_{\nu\nu})\, M(E_{\nu\mu}), \tag{1.7}$$

so daß *alle* $M(E_{\mu\mu})$ positiv sind. Sodann folgt aber weiter aus (1.7), *daß* $M(E_{\mu\nu})$ *für alle* μ *und* ν *positiv ist.*

17. Wenden wir nun IV_M auf (1.5) an, so ergibt sich

$$|a_{\nu\sigma}|\, M(E_{\mu\mu}) \leqq M(E_{\mu\nu})\, M(E_{\sigma\mu})\, M(A). \tag{1.8}$$

Man setze nun

$$p = \operatorname*{Min}_{\mu,\nu} M(E_{\mu\nu}), \qquad P = \operatorname*{Max}_{\mu,\nu} M(E_{\mu\nu}).$$

Wählen wir dann ν und σ in (1.8) so, daß $|a_{\nu\sigma}| = \operatorname*{Max}_{\mu,\nu} |a_{\mu\nu}| = N^*(A)$ ist, so folgt schließlich für $A \neq 0$

$$M(A) \geqq \frac{p}{P^2} N^*(A) > 0,$$

so daß für $M(A)$ auch das Postulat I_N erfüllt ist. Damit ist der Satz I bewiesen.

18. Es sei nun ξ ein Eigenvektor von A, zu einer Fundamentalwurzel λ von A mit $|\lambda| = \lambda_A$ (vgl. Nr. 6) gehörend; dann gilt die Relation $A\xi = \lambda\xi$, die sich in den Bezeichnungen der Nr. 10 auch in der Form $AA_\xi = \lambda A_\xi$ schreiben läßt. Daraus folgt aber wegen II_M und IV_M

$$|\lambda|\, M(A_\xi) \leqq M(A)\, M(A_\xi),$$

und da $M(A_\xi) > 0$ ist, folgt daraus

$$M(A) \geqq \lambda_A, \tag{1.9}$$

d.h. die Behauptung von Satz II.

19. Endlich ergibt sich der Satz III unmittelbar. In der Tat gilt nach (5)

$$\frac{h_M(A\xi)}{h_M(\xi)} = \frac{M(AA_\xi)}{M(A_\xi)} \leqq \frac{M(A)\, M(A_\xi)}{M(A_\xi)} = M(A),$$

und (6) ergibt sich dann ohne weiteres aus der Definition (3).

In der Nr. 9 ist für allgemeine, durch die Formel (2) definierte Schrankennormen die Gültigkeit des Multiplikationspostulates IV_M nur im Falle der symmetrischen Schrankennormen bewiesen worden. Im Falle der allgemeinen

zu zwei Vektornormen h und h_1 gehörenden Schrankennormen gilt indessen eine etwas anders beschaffene „Multiplikationsrelation". Existiert nämlich das Produkt AB und sind h_1, h_2 und h_3 drei Vektornormen, so gilt

$$S_{h_1,h_3}(AB) \leq S_{h_1,h_3}(A)\, S_{h_2,h_3}(B), \tag{1.10}$$

sofern die Dimensionen von h_1, h_2, h_3 bzw. der Zeilenanzahl von A, der Zeilenanzahl von B und der Kolonnenanzahl von B gleich sind. In der Tat ist ja

$$\frac{h_1(AB\xi)}{h_3(\xi)} = \frac{h_1(AB\xi)}{h_2(B\xi)} \cdot \frac{h_2(B\xi)}{h_3(\xi)},$$

woraus (1.10) sofort folgt.

— Im Falle, daß die Normen h_1, h_2, h_3 als Quadratwurzeln aus positiven hermiteschen Formen in den Vektorkomponenten definiert werden, ist die Relation (1.10) von GAUTSCHI ([3], p. 4, Formel (15)) gegeben worden.

§ 2. Allgemeine Normen.

20. Wir beweisen nunmehr den Satz IV. Aus (1.2), II_N und III_N folgt

$$N(A) \leq \sum_{\mu,\nu} |a_{\mu\nu}| N(E_{\mu\nu}) \tag{2.1}$$

und daher

$$|N(A) - N(B)| \leq \sum_{\mu,\nu} |a_{\mu\nu} - b_{\mu\nu}|\, N(E_{\mu\nu}). \tag{2.2}$$

Daher ist $N(A)$ *stetig in den* $a_{\mu\nu}$, und zugleich gilt

$$N(A) \leq N^*(A) \sum_{\mu,\nu} N(E_{\mu\nu}) = p_2 N^*(A), \quad p_2 = \sum_{\mu,\nu} N(E_{\mu\nu}).$$

Andererseits existiert wegen der Stetigkeit von $N(A)$ auf der abgeschlossenen Menge $\operatorname{Max}_{\mu,\nu} |a_{\mu\nu}| = 1$ das

$$\operatorname{Min} N(A) = p_1$$

und ist positiv; daher gilt dann allgemein

$$N(A) \geq p_1 N^*(A),$$

und (7) ist bewiesen.

21. Wir kommen nunmehr zum Beweis des Satzes V. Es möge A durch die Matrix S auf die Jordansche Normalform[9]) entsprechend den Elementarteilern (8) transformiert werden:

$$T = SAS^{-1} = T_1 \dotplus T_2 \dotplus \ldots \dotplus T_k. \tag{2.3}$$

Hierbei *zerfällt* also T *vollständig* in die k quadratischen Matrizen $T_\varkappa$, wo $T_\varkappa$ eine Matrix der Ordnung $m_\varkappa$ ist, die zum Elementarteiler $(\lambda - \lambda_\varkappa)^{m_\varkappa}$ gehört. Jede dieser einzelnen Matrizen $T_\varkappa$ ist gleich

$$\lambda_\varkappa E_{m_\varkappa} + U_{m_\varkappa},$$

[9]) Vgl. z.B. SCHREIER und SPERNER [1], pp. 126—129, sowie namentlich TURNBULL and AITKEN [1], pp. 58—63.

wo $E_{m_\varkappa}$ die Einheitsmatrix der Ordnung $m_\varkappa$ ist, während $U_{m_\varkappa}$ eine quadratische Matrix der Ordnung $m_\varkappa$ ist, in der längs der Hauptdiagonalen und darunter lauter Nullen stehen, in der nächsten Superdiagonalen lauter Einsen und darüber wiederum lauter Nullen. Das Element $u_{\mu\nu}$ von $U_{m_\varkappa}$ ist also gegeben durch das Kroneckersche Symbol $\delta_{\nu,\mu+1}$.

Man beachte, daß wenn eine solche Matrix $U_{m_\varkappa}$ in die Potenz σ $(\sigma = 2, 3, \ldots)$ erhoben wird, dann eine Matrix entsteht, die in der σ-ten Superdiagonalen lauter Einsen hat und sonst durchweg Nullen. Die Elemente $u_{\mu\nu}^{(\sigma)}$ von $U_{m_\varkappa}^{\sigma}$ sind daher gegeben durch $\delta_{\nu,\mu+\sigma}$. Insbesondere enthält $U_{m_\varkappa}^{m_\varkappa-1}$ nur eine Einheit in der rechten oberen Ecke und sonst durchweg Nullen, während $U_{m_\varkappa}^{\sigma}$ für $\sigma \geqq m_\varkappa$ verschwindet.

22. Daher gilt

$$(\lambda_\varkappa E_{m_\varkappa} + U_{m_\varkappa})^t = \lambda_\varkappa^t \left[E_{m_\varkappa} + \binom{t}{1} \lambda_\varkappa^{-1} U_{m_\varkappa} + \binom{t}{2} \lambda_\varkappa^{-2} U_{m_\varkappa}^2 + \cdots + \binom{t}{m_\varkappa - 1} \lambda_\varkappa^{-m_\varkappa+1} U_{m_\varkappa}^{m_\varkappa-1}\right].$$

Daraus folgt für den maximalen absoluten Betrag eines Elementes der Matrix links die für hinreichend große t geltende Relation

$$N^*(T_\varkappa^t) = \binom{t}{m_\varkappa - 1} |\lambda_\varkappa|^{t-m_\varkappa+1}.$$

Darüber hinaus gilt für $t \to \infty$ offenbar

$$T_\varkappa^t = \binom{t}{m_\varkappa - 1} \lambda_\varkappa^{t-m_\varkappa+1} E_{1m_\varkappa}^{(\varkappa)} + O(t^{m_\varkappa-2} |\lambda_\varkappa|^t),$$

wo $E_{1m_\varkappa}^{(\varkappa)}$ eine quadratische Matrix der Ordnung $m_\varkappa$ ist, die in der rechten oberen Ecke eine Eins enthält und sonst lauter Nullen.

23. Nunmehr folgt aus (2.3) für hinreichend große t

$$N^*(T^t) = \operatorname*{Max}_\varkappa \binom{t}{m_\varkappa - 1} |\lambda_\varkappa|^{t-m_\varkappa+1} = \lambda_A^t \operatorname*{Max}_{\varkappa \leqq s} \binom{t}{m_\varkappa - 1} \lambda_A^{-m_\varkappa-1} = \binom{t}{M-1} \lambda_A^{t-M+1},$$

$$N^*(T^t) = \binom{t}{M-1} \lambda_A^{t-M+1} \tag{2.4}$$

und, $\lambda_\nu/\lambda_A = \varepsilon_\nu$ $(\nu = 1, 2, \ldots, s)$ gesetzt,

$$T^t = \binom{t}{M-1} \lambda_A^{t-M+1} [E_{1M} \varepsilon_1^{t-M+1} + E_{M+1,2M} \varepsilon_2^{t-M+1} + \cdots + E_{(s-1)M+1,sM} \varepsilon_s^{t-M+1}] + O(t^{M-2} \lambda_A^t),$$

wo nunmehr die Symbole $E_{\mu\nu}$ die in der Nr. 15 festgelegte Bedeutung haben. Daraus folgt im Falle, daß die Relation (12) gilt, da ja dann alle ε_ν einander gleich sind,

$$T^t = \binom{t}{M-1} \lambda_1^{t-M+1} \sum_{\sigma=1}^{s} E_{(\sigma-1)M+1,\sigma M} + O(t^{M-2} \lambda_A^t). \tag{2.5}$$

24. Wegen (2.3) gilt $T^t = SA^tS^{-1}$, $A^t = S^{-1}T^tS$ für jedes natürliche t und daraus folgt für

$$\text{(2.6)} \qquad \begin{cases} S = (s_{\mu\nu}), \quad S^{-1} = (s'_{\mu\nu}), \quad \sigma = \sum_{\mu,\nu} |s_{\mu\nu}| \sum_{\mu,\nu} |s'_{\mu\nu}| : \\ N^*(T^t) \leqq \sigma N^*(A^t), \quad N^*(A^t) \leqq \sigma N^*(T^t) \end{cases}$$

und daher wegen (2.4)

$$\text{(2.7)} \qquad \frac{1}{\sigma} \leqq \frac{N^*(A^t)}{\binom{t}{M-1}\lambda_A^{t-M+1}} \leqq \sigma,$$

womit wegen (7) auch (11) bewiesen ist.

25. Andererseits folgt aus (2.5) für $t \to \infty$

$$\frac{A}{\binom{t}{M-1}\lambda_1^{t-M+1}} = C + O\left(\frac{1}{t}\right), \qquad C = \sum_{\sigma=1}^{s} S^{-1}E_{(\sigma-1)M+1,\sigma M}S.$$

Daher gilt wegen II_N in Nr. 3 für $t \to \infty$

$$\frac{N(A^t)}{\binom{t}{M-1}\lambda_A^{t-M+1}} \to N(C).$$

Wäre nun $N(C) = 0$, so wäre $N(A^t) = o(t^{M-1}\lambda_A^t)$, entgegen (2.7). Daher folgt die zweite Behauptung des Satzes V unmittelbar[10]).

§ 3. Höldersche Normen.

26. Von nun an sollen die Matrizen $A, B, \ldots$ nur als rechteckige Matrizen vorausgesetzt werden. Auch für solche Matrizen, z.B. Matrizen vom $(m \times n)$-Typus, d.h. mit m Zeilen und n Kolonnen, soll der Normbegriff durch die Postulate I_N bis III_N der Nr. 3 definiert werden. Eine solche Norm kann z.B. in sehr allgemeiner Weise aus gegebenen Vektornormen für n- und m-dimensionale Vektoren gebildet werden. Es seien $h_1(\xi), \ldots, h_n(\xi)$ n gleiche oder verschiedene Vektornormen für m-dimensionale Vektoren ξ und es sei $h(\eta)$ eine Vektornorm für n-dimensionale Vektoren η. $h(\eta)$ soll ferner die Eigenschaft haben, daß sie für nicht negative Komponenten $y_1, \ldots, y_n$ von η eine nicht abnehmende Funktion dieser Komponenten ist. Bezeichnen wir dann mit ξ_ν den ν-ten Kolonnenvektor der Matrix A, so daß also

$$\xi_\nu = (a_{1\nu}, a_{2\nu}, \ldots, a_{m\nu})$$

gesetzt, so setze man

$$\text{(3.1)} \qquad H(A) = h[h_1(\xi_1), h_2(\xi_2), \ldots, h_n(\xi_n)].$$

Damit ist eine Norm von A definiert.

[10]) Wir hätten anstatt der Jordanschen Normalform auch z.B. die explizite Darstellung von A, wie sie sich bei H. Schmidt ([1], p. 549, Formel (9)) findet, benutzen können, doch wäre hierzu erst eine weitere Diskussion der in diese Formel eingehenden Vektoren notwendig.

27. In der Tat, daß die Postulate I_N und II_N der Nr. 3 erfüllt sind, ist unmittelbar klar. Bei III_N ergibt sich dies aber wie folgt: Es sei A^0 eine weitere $(m \times n)$-Matrix und es seien ihre Kolonnenvektoren $\xi_1^0, \ldots, \xi_n^0$. Dann ergibt sich wegen (3.1) unter Benutzung der Postulate der Nr. 7:

$$H(A + A^0) = h[h_1(\xi_1 + \xi_1^0), \ldots, h_n(\xi_n + \xi_n^0)] \leq$$
$$\leq h[h_1(\xi_1) + h_1(\xi_1^0), \ldots, h_n(\xi_n) + h_n(\xi_n^0)] \leq$$
$$\leq h[h_1(\xi_1), \ldots, h_n(\xi_n)] + h[h_1(\xi_1^0), \ldots, h_n(\xi_n^0)] = H(A) + H(A^0).$$

28. Definieren wir z. B. allgemein für einen k-dimensionalen Vektor ξ und für ein p, $1 \leq p \leq \infty$,

$$|\xi|_p = \left(\sum_\varkappa |x_\varkappa|^p\right)^{1/p} \qquad (1 \leq p \leq \infty), \tag{3.2}$$

wobei natürlich für $p = \infty$, $|\xi|_\infty = \operatorname{Max}_\varkappa |x_\varkappa|$ ist, so ist $|\xi|_p$ eine Norm im Sinne der Postulate der Nr. 7. Dabei folgt insbesondere das Postulat III_V aus der Minkowskischen Ungleichung.

Setzen wir nun insbesondere in (3.1) $h_\nu(\xi) = |\xi|_{p_1}$ und $h(\eta) = |\eta|_{p_2}$, so erhalten wir die Norm

$$|A|_{p_1, p_2} = \left[\sum_\nu \left(\sum_\mu |a_{\mu\nu}|^{p_1}\right)^{p_2/p_1}\right]^{1/p_2}, \tag{3.3}$$

wo p_1 und p_2 zwischen 1 und ∞ liegen. Andererseits ergibt sich, wenn (3.2) auf den Fall $k = m\,n$ angewandt wird, daß der für eine Matrix A durch (14) definierte Ausdruck für $1 \leq p \leq \infty$ eine Norm ist.

Man beachte, daß während $|A|_p$ beim Übergang zur transponierten Matrix unverändert bleibt, $|A'|_p = |A|_p$, dies bei der Norm (3.3) im allgemeinen nicht der Fall ist. — Ist die Kolonnenanzahl in A gleich 1, so gilt, wie man sich sofort überzeugt,

$$|A|_{p_1, p_2} = |A|_{p_1}.$$

29. Mit Hilfe von (3.2) läßt sich die Höldersche Ungleichung so formulieren: Gilt

$$\frac{1}{p} + \frac{1}{q} = 1, \qquad 1 \leq p \leq \infty, \tag{3.4}$$

so gilt für das innere Produkt

$$(\xi, \eta) = \sum_{\nu=1}^{n} x_\nu \bar{y}_\nu \tag{3.5}$$

der Vektoren ξ, η der Dimension n, die Relation

$$|(\xi, \eta)| \leq |\xi|_p |\eta|_q. \tag{3.6}$$

Andererseits gilt bekanntlich, daß der Ausdruck (3.2) *mit zunehmendem* $p \geq 1$ *monoton abnimmt.* (Diese Tatsache ist zuerst von A. PRINGSHEIM angegeben worden (vgl. HARDY-LITTLEWOOD-PÓLYA [1], p. 28).)

Ist nun $p \leq 2$, so ist wegen (3.4) $q \geq 2$, $p \leq q$, so daß

$$|\xi|_q \leq |\xi|_p \qquad (p \leq 2 \leq q) \tag{3.7}$$

gilt. Daher folgt nunmehr aus (3.6), wenn (3.7) auf η angewendet wird,

$$|(\xi,\eta)| \leq |\xi|_p |\eta|_p \qquad (1 \leq p \leq 2). \tag{3.8}$$

30. Wir beweisen nun einen Hilfssatz über den Ausdruck (3.3) unter der Annahme, daß $p_1 = q$ und $p_2 = p$ durch die Relation (3.4) verknüpft sind.

Hilfssatz. *Gilt* (3.4), *so gilt für zwei Matrizen* A, B, *für die das Produkt* AB *existiert*:

$$|AB|_{q,p} \leq |A|_q |B|_p. \tag{3.9}$$

Ist aber darüber hinaus $p \geq 2$, *so gilt*

$$|A|_{q,p} \leq |A|_q \qquad (p \geq 2). \tag{3.10}$$

31. *Beweis.* Es sei A eine $(m \times k)$- und B eine $(k \times n)$-Matrix. Dann ist $AB = C = (c_{\mu\nu})$ eine $(m \times n)$-Matrix und es gilt wegen $c_{\mu\nu} = \sum_{\varkappa} a_{\mu\varkappa} b_{\varkappa\nu}$ nach (3.6), wenn dort p mit q vertauscht wird und beide Seiten in die q-te Potenz erhoben werden,

$$|c_{\mu\nu}|^q \leq \Big(\sum_{\varkappa} |a_{\mu\varkappa}|^q\Big)\Big(\sum_{\varkappa} |b_{\varkappa\nu}|^p\Big)^{q/p}$$

und daher, wenn nach μ summiert wird, wegen (14),

$$\sum_{\mu} |c_{\mu\nu}|^q \leq |A|_q^q \Big(\sum_{\varkappa} |b_{\varkappa\nu}|^p\Big)^{q/p},$$

$$\Big(\sum_{\mu} |c_{\mu\nu}|^q\Big)^{p/q} \leq |A|_q^p \sum_{\varkappa} |b_{\varkappa\nu}|^p,$$

und wenn hier über ν summiert wird, wegen (3.3) und (14),

$$|AB|_{q,p}^p \leq |A|_q^p |B|_p^p,$$

womit (3.9) bewiesen ist.

32. Andererseits läßt sich, wenn

$$y_\nu = \Big(\sum_{\mu} |a_{\mu\nu}|^q\Big)^{1/q},$$

gesetzt wird, $|A|_{q,p}$ schreiben als

$$|A|_{q,p} = \Big(\sum_{\nu} y_\nu^p\Big)^{1/p}$$

und dies ist, wenn $p \geq 2 \geq q$ ist, nach (3.7), wenn dort p mit q vertauscht wird, höchstens gleich

$$\Big(\sum_{\nu} y_\nu^q\Big)^{1/q} = \Big(\sum_{\nu,\mu} |a_{\mu\nu}|^q\Big)^{1/q} = A_q,$$

womit (3.10) bewiesen ist.

33. Um nun den Satz VII zu beweisen, nehmen wir an, daß A eine $(m \times k)$- und B eine $(k \times n)$-Matrix ist. Ist dann $c_{\mu\nu}$ das allgemeine Element von $C = AB$, so gilt $c_{\mu\nu} = \sum_{\varkappa} a_{\mu\varkappa} b_{\varkappa\nu}$ und daher wegen (3.8) für $p \leq 2$

$$|c_{\mu\nu}|^p \leq \sum_{\varkappa} |a_{\mu\varkappa}|^p \sum_{\varkappa} |b_{\varkappa\nu}|^p.$$

Summieren wir hier über μ und ν, so folgt

$$|C|_p^p \leq |A|_p^p |B|_p^p,$$

woraus (15) unmittelbar folgt.

34. Ist aber $p > 2$, so wählen wir als Matrizen A und B die Matrizen deren sämtliche Elemente gleich 1 sind. Dann ist $c_{\mu\nu} = k$ und

$$|A|_p = (m\,k)^{1/p}, \quad |B|_p = (n\,k)^{1/p}, \quad C_p = k\,(m\,n)^{1/p}.$$

Hier liefe die Relation (15) hinaus auf

$$(m\,n)^{1/p}\,k \leq m^{1/p}\,k^{1/p}\,n^{1/p}\,k^{1/p},$$

was für $p > 2$ unmöglich ist, solange $k > 1$ ist. Für $k = 1$ dagegen ist (15) offenbar für alle $p \geq 1$ richtig, sogar mit dem Gleichheitszeichen. Damit ist der Satz VII bewiesen.

35. Es sei wiederum A eine $(m \times k)$- und B eine $(k \times n)$-Matrix. Dann gilt nach (3.6) für $c_{\mu\nu} = \sum_{\varkappa} a_{\mu\varkappa} b_{\varkappa\nu}$:

$$|c_{\mu\nu}|^p \leq \sum_{\varkappa} |a_{\mu\varkappa}|^p \left(\sum_{\varkappa} |b_{\varkappa\nu}|^q\right)^{p/q}$$

und daher, wenn nach μ und ν summiert wird, wegen (3.3)

$$|AB|_p^p \leq |A|_p^p \sum_{\nu} \left(\sum_{\varkappa} |b_{\varkappa\nu}|^q\right)^{p/q} = |A|_p^p |B|_{q,p}^p,$$

$$|AB|_p \leq |A|_p |B|_{q,p} \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right). \tag{3.11}$$

Andererseits folgt wiederum aus (3.6), wenn dort p mit q vertauscht wird

$$|c_{\mu\nu}|^p \leq \sum_{\varkappa} |b_{\varkappa\nu}|^p \left(\sum_{\varkappa} |a_{\mu\varkappa}|^q\right)^{p/q},$$

und daraus durch Summation nach μ und ν

$$|AB|_p^p \leq |B|_p^p \sum_{\mu} \left(\sum_{\varkappa} |a_{\mu\varkappa}|^q\right)^{p/q}.$$

Der zweite Faktor rechts ist aber gleich $|A'|_{q,p}^p$, unter A' die zu A gehörende transponierte Matrix verstanden, so daß wir schließlich erhalten

$$|AB|_p \leq |B|_p |A'|_{q,p}. \tag{3.12}$$

36. Ist nun $p \geq 2$, so folgt aus (3.11) und (3.10), auf B angewandt, die Relation (16). Aus (3.12) und (3.10) folgt aber

$$|AB|_p \leq |B|_p |A'|_q = |B|_p |A|_q,$$

womit auch (17) bewiesen ist.

37. Ist jetzt $p < 2$, so spezialisiere man A und B wiederum zu Matrizen deren sämtliche Elemente gleich 1 sind. Dann gelten die Relationen

$$|A|_p = (m\,k)^{1/p}, \quad |B|_q = (n\,k)^{1/q}, \quad |C|_p = k\,(m\,n)^{1/p}.$$

Dann liefe die Ungleichung (16) hinaus auf

$$k\,(m\,n)^{1/p} \leq m^{1/p}\,k^{1/p}\,n^{1/q}\,k^{1/q},$$

d.h. wegen (3.4) auf $n^{1/p} \leqq n^{1/q}$, was für $n > 1$, $p < 2$ unmöglich ist. Ebenso ist (17) für $m > 1$, $p < 2$ unmöglich.

38. Ist aber $n = 1$, so folgt aus der Definition (3.3), daß $|B|_{q,p} = |B|_q$, so daß (16) für alle $p \geqq 1$ aus (3.11) folgt. Ebenso reduziert sich für $m = 1$, $|A'|_{q,p}$ auf $|A'|_q = |A|_q$, so daß dann (17) aus (3.12) für alle $p \geqq 1$ folgt. Damit ist der Satz VIII bewiesen.

39. Um die Relation (18) des Satzes IX zu beweisen, beachte man, daß aus (3.11) folgt

$$|A_1 A_2 \dots A_u|_p \leqq |A_1|_p \, |A_2 \dots A_u|_{q,p} \tag{3.13}$$

und daher, wenn auf den zweiten Faktor rechts (3.9) angewendet wird,

$$|A_1 A_2 \dots A_u|_p \leqq |A_1|_p \, |A_2|_q \, |A_3 \dots A_u|_p, \tag{3.14}$$

woraus (18) durch vollständige Induktion hervorgeht.

40. Soll nun die Relation (19) allgemein gelten und ist A_1 eine $(m \times k_1)$-, A_2 eine $(k_1 \times k_2)$-, ..., A_g eine $(k_{g-1} \times n)$-Matrix, so spezialisiere man alle diese Matrizen in der Weise, daß ihre sämtlichen Elemente gleich 1 gesetzt werden. Dann ist das Produkt eine $(m \times n)$-Matrix, deren sämtliche Elemente gleich $k_1 k_2 \dots k_{g-1}$ sind. Daher wird die linke Seite von (19) gleich

$$(m\,n)^{1/p} \, k_1 \, k_2 \dots k_{g-1},$$

während das Produkt rechts in (19) zu

$$(m\,k_1)^{1/p} \, (k_1 \, k_2)^{1/q} \dots (k_{g-2} \, k_{g-1})^{1/p} \, (k_{g-1} \, n)^{1/q} = m^{1/p} \, n^{1/q} \, k_1 \dots k_{g-1}$$

wird. Aus (19) folgt aber dann $n^{1/p} \leqq n^{1/q}$, was nur für $n = 1$ oder $p \geqq 2 \geqq q$ gelten kann.

Um aber einzusehen, daß in den beiden letzten Fällen (19) gilt, beachte man, daß im Falle eines geraden g genau wie oben aus (3.13) und (3.14) durch vollständige Induktion die Formel

$$|A_1 A_2 \dots A_g|_p \leqq |A_1|_p \, |A_2|_q \dots |A_{g-1}|_p \, |A_g|_{q,p} \tag{3.15}$$

folgt. Für $p \geqq 2$ folgt aber dann (19) aus (3.10) und für $n = 1$ daraus, daß dann wegen (3.3) der letzte Faktor in (3.15) gleich $|A_g|_q$ wird. Damit ist Satz IX bewiesen.

41. Soll nun die Relation (20) des Satzes X allgemein gelten, so spezialisiere man die Matrizen $A_1, \dots, A_l$ wiederum derart, daß ihre sämtlichen Elemente gleich 1 werden, und nehme die zugehörigen Zeilen- und Kolonnenanzahlen, wie in der Nr. 40, bzw. gleich $m, k_1, k_2, \dots, k_{l-1}, n$. Dann ist $A_1 \dots A_l$ eine $(m \times n)$-Matrix, deren sämtliche Elemente gleich $k_1 \dots k_{l-1}$ sind, so daß der Ausdruck links in (20) zu $(m\,n)^{1/p} k_1 \dots k_{l-1}$ wird. Das Produkt rechts in (20) wird im Falle eines ungeraden l, wie in der Nr. 40, zu $(m\,k_1)^{1/q} (k_1 k_2)^{1/p} \dots (k_{l-1} n)^{1/q}$, so daß hier (20) auf

$$(m\,n)^{1/p} \leqq (m\,n)^{1/q}$$

hinausläuft, und wir entweder $m=n=1$ oder $p\geqq 2\geqq q$ erhalten. Für ein gerades l dagegen wird das Produkt rechts in (20) zu $(m\,k_1)^{1/q}(k_1k_2)^{1/p}\dots(k_{l-1}n)^{1/p}$ und (20) läuft auf

$$(m\,n)^{1/p}=m^{1/q}\,n^{1/p}$$

hinaus, so daß entweder $m=1$ oder $p\geqq 2\geqq q$ ist.

42. Ist andererseits $p\geqq 2$, so folgt aus (17)

$$|A_1\dots A_l|_p\leqq|A_1|_q\,|A_2\dots A_l|_p$$

und (20) bzw. (20°) folgt hier aus (18) oder (19), da ja auch die Relation (19) für $p\geqq 2$ gilt. Ist aber $m=1$, so folgt aus (3.12)

$$|A_1\dots A_l|_p\leqq|A_2\dots A_l|_p\,|A_1'|_{q,p}.$$

Der letzte Faktor ist aber gleich $|A_1'|_q=|A_1|_q$, da die Kolonnenzahl m in A_1' den Wert 1 hat. Daher folgt

$$|A_1\dots A_l|_p\leqq|A_1|_q\,|A_2\dots A_l|_p\qquad(m=1).\tag{3.16}$$

Ist nun g gerade, also die Faktorenzahl in $A_2\dots A_g$ ungerade, so folgt (20°) aus (3.16) wegen (18). Ist dagegen l ungerade, so haben wir auch zugleich $n=1$ vorausgesetzt. In diesem Falle ist aber auf den zweiten Faktor rechts in (3.16) die Ungleichung (19) anzuwenden, so daß (20) folgt. Damit ist der Satz X bewiesen.

43. Zum Beweise des Satzes XI genügt es offenbar (23) zu beweisen, da (22) daraus für $A_1=A_2=\dots=A_l$, $k=1$ ohne weiteres folgt. Nun folgt aber aus (3.12)

$$|A_1\dots A_l|_p\leqq|A_1'|_{q,p}\,|A_2\dots A_l|_p\leqq\dots\leqq|A_1'|_{q,p}\dots|A_{k-1}'|_{q,p}\,|A_kA_{k+1}\dots A_l|_p.$$

Für den letzten Faktor rechts folgt durch wiederholte Anwendung von (3.11)

$$|A_k\dots A_l|_p\leqq|A_k\dots A_{l-1}|_p\,|A_l|_{q,p}\leqq\dots\leqq|A_k\dots A_{l-2}|_p\,|A_{l-1}|_{q,p}\,|A_l|_{q,p}\leqq\dots\leqq$$
$$\leqq|A_k|_p\,|A_{k+1}|_{q,p}\dots|A_l|_{q,p}.$$

Damit ist Satz XI bewiesen.

44. Wir wollen noch einige Angaben über diejenigen Schrankennormen [vgl. die Definition in (2)] hinzufügen, die zu den durch (3.2) definierten Hölderschen Vektornormen gehören. Wir setzen

$$S_{p_1,p_2}(A)=\operatorname{Sup}\frac{|A\xi|_{p_1}}{|\xi|_{p_2}}\qquad(1\leqq p_1,p_2\leqq\infty)\tag{3.17}$$

und ferner

$$S_p(A)=S_{p,p}(A)\qquad(1\leqq p\leqq\infty).\tag{3.18}$$

Aus dem in Nr. 29 erwähnten monotonen Abnehmen von (3.2) folgt, daß $S_{p_1,p_2}(A)$ eine *monoton abnehmende* Funktion von p_1 und eine *monoton zunehmende* Funktion von p_2 ist. Von den symmetrischen Schrankennormen

(3.18) lassen sich namentlich die $p=1$ und $p=\infty$ entsprechenden sehr einfach darstellen:

$$S_1(A) = \operatorname{Max}_{\nu} \sum_{\mu} |a_{\mu\nu}|, \tag{3.19}$$

$$S_\infty(A) = \operatorname{Max}_{\mu} \sum_{\nu} |a_{\mu\nu}|. \tag{3.20}$$

Diese beiden sehr leicht zu beweisenden Formeln werden häufig in der Iterationstheorie der Matrizen verwendet[11]).

45. Es ist von Interesse, den Wert der Norm (3.17) für $A=E$, wo E eine n-dimensionale Einheitsmatrix ist, zu bestimmen. Hierzu ist es offenbar notwendig das Extremalproblem zu lösen

$$\operatorname{Max}\left(\sum_{\nu=1}^{n} x_\nu^{p_1}\right)^{1/p_1} = ? \qquad \left(\sum_{\nu=1}^{n} x_\nu^{p_2} = 1\right). \tag{3.21}$$

Ist $p_1 \geqq p_2$, so folgt aus der oben erwähnten Monotonieeigenschaft von $S_{p_1,p_2}(A)$, daß $S_{p_1,p_2}(E) \leqq S_{p_2,p_2}(E) = 1$ ist, und die Schranke 1 wird erreicht für $x_1 = 1$, $x_2 = x_3 = \cdots = x_n = 0$. Für $p_1 < p_2$ liefert die Lagrangesche Multiplikatorenmethode sofort das Resultat, daß wenn keines der x_ν im Extremalpunkt verschwindet, $x_1 = x_2 = \cdots = x_n = n^{-1/p_2}$ und dafür $|\xi|_{p_1} = n^{1/p_1 - 1/p_2} > 1$ ist. Wären aber einige der $x_\nu = 0$, so hätte man hier n durch eine kleinere Zahl zu ersetzen, so daß sich schließlich ergibt:

$$S_{p_1,p_2}(E) = \begin{cases} 1 & (p_1 \geqq p_2) \\ n^{\frac{1}{p_1} - \frac{1}{p_2}} & (p_1 \leqq p_2). \end{cases} \tag{3.22}$$

46. Existiert das Matrizenprodukt AB, so folgt, wenn (18) auf das Zählerprodukt angewandt wird:

$$S_{p_1,p_2}(AB) = \operatorname{Sup} \frac{|AB\xi|_{p_1}}{|\xi|_{p_2}} \leqq |A|_{p_1} |B|_{q_1} S_{p_1,p_2}(E) \qquad \left(\frac{1}{p_1} + \frac{1}{q_1} = 1\right)$$

und daher insbesondere wegen (3.22)

$$S_{p_1,p_2}(AB) \leqq |A|_{p_1} |B|_{q_1} \qquad (p_1 \geqq p_2). \tag{3.24}$$

Ferner folgt, wenn die Sätze VII und VIII auf $|A\xi|_{p_1}$ angewandt werden,

$$S_{p_1,p_2}(A) \leqq |A|_{p_1} S_{p_1,p_2}(E) \qquad (p_1 \leqq 2), \tag{3.25}$$

$$S_{p_1,p_2}(A) \leqq |A|_{p_1} S_{q_1,p_2}(E) \qquad (1 \leqq p_1 \leqq \infty), \tag{3.26}$$

$$S_{p_1,p_2}(A) \leqq |A|_{q_1} S_{p_1,p_2}(E) \qquad (p_1 \geqq 2). \tag{3.27}$$

Es sei endlich noch vermerkt, daß

$$S_{\infty,1}(A) = N^*(A) = \operatorname{Max}_{\mu,\nu} |a_{\mu\nu}| \tag{3.28}$$

[11]) Vgl. z.B. Faddejewa [1], pp. 62, 63; NBSR pp. 85—87.

ist. In der Tat gilt

$$\frac{\operatorname*{Max}_{\mu} \left| \sum_{\nu} a_{\mu\nu} x_\nu \right|}{\sum_{\nu} |x_\nu|} \leq \operatorname*{Max}_{\mu,\nu} |a_{\mu\nu}|,$$

und die Schranke rechts wird offenbar erreicht, wenn ein geeignetes der x_ν gleich 1 und alle übrigen gleich 0 gesetzt werden.

Literatur.

AITKEN, A. C., and H. W. TURNBULL: [*1*] An Introduction to the Theory of Canonical Matrices. London and Glasgow, Blackie and Son 1950. — FADDEJEWA, W. N.: [*1*] Numerische Methoden der linearen Algebra. Moskau 1950 (russisch). Eine Übersetzung des für uns vor allem in Frage kommenden 1. Kapitels des Buches ist veröffentlicht worden als National Bureau of Standards Report 1644, Computational methods of linear algebra, Ch. 1, Basic material from linear algebra, pp. V + 1—93 (Translator: C. D. Benster, Editor: G. E. Forsythe). Im Text werden wir die Zitate auf diese Übersetzung mit NBSR charakterisieren. — GAUTSCHI, WE.: [*1*] The asymptotic behaviour of powers of matrices. Duke Math. J. **20**, 127—140 (1953). — [*2*] The asymptotic behaviour of powers of matrices. II. Duke Math. J. **20**, 375—379 (1953). — [*3*] Bounds of matrices with regard to an Hermitian metric. Comp. Math. **12**, 1—16 (1954). — HARDY, LITTLEWOOD, PÓLYA: [*1*] Inequalities. Cambridge 1934. — OSTROWSKI, A.: [*1*] Sur la détermination des bornes inférieures pour une classe des déterminants. Bull. Sci. Math. (2) **61**, 1—14 (1937). — [*2*] Un nouveau théorème d'existence pour les systèmes d'équations. C. R. Acad. Sci. Paris **232**, 786—788 (1951). — RIESZ, F., et B. NAGY: [*1*] Lecons d'analyse fonctionelle, deuxième edition. Budapest 1953. — SCHMIDT, H.: [*1*] Über Wurzelapproximation nach Euler und Fixgebilde linearer Transformationen. Math. Z. **52**, 547—556 (1950). — SCHREIER, O., u. E. SPERNER: [*1*] Einführung in die analytische Geometrie und Algebra, Bd. II. Leipzig 1935.

Mathematische Anstalt der Universität Basel (Schweiz).

(Eingegangen am 24. Dezember 1954.)

Perron, O.
Math. Zeitschr. Bd. 63, S. 19—30 (1955)

Über die Abhängigkeit von Potenzsummen und einen Satz von Pólya.

Issai Schur zum Gedächtnis.

Von

Oskar Perron.

Zu den Lieblingsgebieten von I. Schur haben die algebraischen Gleichungen und die Eigenschaften ihrer Wurzeln gehört. Noch in der letzten Arbeit, deren Erscheinen in Band 4 der Compositio Mathematica (1937) er erleben durfte, hat er sich mit den Potenzsummen befaßt. Wenn er auch zumeist wohl arithmetische Fragen im Auge hatte, so dürften doch die nachstehenden Betrachtungen, bei denen es sich um Potenzsummen von *n Variablen (Unbestimmten)* handelt, in einem Band, der seinem Andenken gewidmet ist, nicht fehl am Ort sein.

§ 1. Abhängigkeit von Potenzsummen.

Zwischen $n+1$ Formen (d.h. homogenen Polynomen) $f_0, f_1, \ldots, f_n$ von n Variablen $x_1, \ldots, x_n$ besteht stets eine Abhängigkeit

$$\sum A_{\lambda_0 \lambda_1 \ldots \lambda_n} f_0^{\lambda_0} f_1^{\lambda_1} \cdots f_n^{\lambda_n} = 0,$$

und zwar, wenn mit m_ν der Grad von f_ν und mit d der größte gemeinsame Teiler von $m_0, m_1, \ldots, m_n$ bezeichnet wird, eine Abhängigkeit von der Höhe (dem Gewicht)

$$H = \frac{m_0 m_1 \cdots m_n}{d},$$

d.h. jedes Glied der Summe ist in bezug auf die x_ν vom Grad

$$m_0 \lambda_0 + m_1 \lambda_1 + \cdots + m_n \lambda_n = H.$$

Wenn die f_ν *allgemeine* Formen sind, d.h., wenn die Koeffizienten Unbestimmte sind, dann besteht zudem keine Abhängigkeit von kleinerer Höhe[1]). Zwischen spezialisierten Formen kann aber sehr wohl eine Abhängigkeit kleinerer Höhe bestehen. Im folgenden sollen speziell gewisse symmetrische Formen betrachtet werden.

[1]) Dieser Doppelsatz ist bewiesen in meiner Arbeit: „Über die Abhängigkeit von Polynomen", Sitzungsberichte Bayer. Akad. Wiss., Math.-nat. Klasse **1950**, 117—130. Einen zweiten Beweis hat Herr Heinz Orsinger in seiner Dissertation gegeben: „Zur Konstruktion von Trägheitsformen als Koeffizienten algebraischer Gleichungen." Math. Nachr. **5**, 355—370 (1951).

Trivial ist zunächst folgendes: Sind

$$\begin{aligned} a_1 &= -(x_1 + x_2 + \cdots + x_n), \\ a_2 &= +(x_1 x_2 + x_1 x_3 + \cdots + x_{n-1} x_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_n &= (-1)^n x_1 x_2 \cdots x_n \end{aligned}$$

die symmetrischen Grundfunktionen und ist f eine symmetrische Form vom Grad m, so ist f bekanntlich als Polynom von $a_1, a_2, \ldots, a_n$ darstellbar:

$$f = \sum B_{\lambda_1 \lambda_2 \ldots \lambda_n} a_1^{\lambda_1} a_2^{\lambda_2} \ldots a_n^{\lambda_n},$$

wobei jedes Glied der Summe vom Gewicht $\lambda_1 + 2\lambda_2 + \cdots + n\lambda_n = m$ ist. Diese Formel bedeutet bereits eine Abhängigkeit der Höhe m zwischen $a_1, \ldots, a_n$ und f, und eine Abhängigkeit von kleinerer Höhe besteht offenbar nicht. Insbesondere kann man die Potenzsummen

$$s_\nu = x_1^\nu + x_2^\nu + \cdots + x_n^\nu$$

so darstellen, und wenn etwa $\nu < n$ ist, s_ν schon als Polynom von $a_1, \ldots, a_\nu$. Die Darstellung ergibt sich leicht in Determinantenform aus den bekannten NEWTONschen Formeln, die ebenso auch umgekehrt a_ν als Polynom von $s_1, \ldots, s_\nu$ darzustellen gestatten. Infolgedessen ist jede symmetrische Form f vom Grad m auch in der Gestalt

$$f = \sum C_{\lambda_1 \lambda_2 \ldots \lambda_n} s_1^{\lambda_1} s_2^{\lambda_2} \ldots s_n^{\lambda_n}$$

darstellbar, wobei offenbar wieder $\lambda_1 + 2\lambda_2 + \cdots + n\lambda_n = m$ ist. Diese Formel bedeutet eine Abhängigkeit der Höhe m zwischen $s_1, \ldots, s_n$ und f, und eine Abhängigkeit kleinerer Höhe besteht wieder nicht. Insbesondere die Abhängigkeit zwischen $s_1, \ldots, s_n, s_{n+1}$ ergibt sich, indem man $a_1, \ldots, a_n$ aus den ersten $n+1$ NEWTONschen Formeln eliminiert, in der Gestalt:

$$\begin{vmatrix} s_1 & 1 & & \\ s_2 & s_1 & 2 & \\ \cdots & \cdots & \cdots & \cdots \\ s_n & s_{n-1} & s_{n-2} \cdots s_1 & n \\ s_{n+1} & s_n & s_{n-1} \cdots s_2 & s_1 \end{vmatrix} = 0,$$

wo die leeren Stellen durch Nullen zu besetzen sind.

Ganz anders und durchaus nicht mehr trivial ist die Sache, wenn nur Potenzsummen mit ungeraden Exponenten betrachtet werden. Die ersten n dieser Potenzsummen $s_1, s_3, s_5, \ldots, s_{2n-1}$ sind augenscheinlich voneinander unabhängig, denn ihre Funktionaldeterminante ist eine nicht verschwindende VANDERMONDEsche Determinante. Dagegen besteht zwischen $s_1, s_3, s_5, \ldots, s_{2n+1}$ nach obigem Satz eine Abhängigkeit der Höhe $1 \cdot 3 \cdot 5 \cdots (2n+1)$. Aber das ist nicht die kleinstmögliche Höhe, vielmehr werden wir folgendes beweisen:

Satz 1. *Zwischen den $n+1$ Potenzsummen*

$$s_\nu = x_1^\nu + x_2^\nu + \cdots + x_n^\nu \qquad (\nu = 1, 3, 5, \ldots, 2n+1)$$

besteht genau eine Abhängigkeit der Höhe $H = (n+1)(n+2)/2$, die zudem in bezug auf s_{2n+1} linear ist, so daß s_{2n+1} sich rational durch $s_1, s_3, \ldots, s_{2n-1}$ ausdrücken läßt. Es besteht aber keine Abhängigkeit von kleinerer Höhe.

Außerdem werden wir die Abhängigkeit der Höhe H auch formelmäßig angeben können.

Zunächst ist klar, daß es *höchstens eine* Abhängigkeit der Höhe H geben kann. Denn gäbe es zwei verschiedene, so könnte man daraus s_{2n+1} eliminieren und hätte eine Abhängigkeit zwischen $s_1, s_3, \ldots, s_{2n-1}$, während diese n Formen doch voneinander unabhängig sind. Nunmehr zeigen wir, daß eine Abhängigkeit von kleinerer Höhe als H

$$\text{(I)} \qquad \begin{cases} \sum D_{\lambda_1 \lambda_2 \ldots \lambda_{2n+1}} s_1^{\lambda_1} s_3^{\lambda_3} \cdots s_{2n+1}^{\lambda_{2n+1}} = 0, \\ \lambda_1 + 3\lambda_3 + 5\lambda_5 + \cdots + (2n+1)\lambda_{2n+1} = H' < H \end{cases}$$

jedenfalls nicht besteht. Nehmen wir nämlich an, sie bestünde doch, so bilden wir nach Einführung von zwei weiteren Variablen x_{n+1}, x_{n+2} die Potenzsummen

$$S_\nu = x_1^\nu + x_2^\nu + \cdots + x_{n+1}^\nu + x_{n+2}^\nu = s_\nu + x_{n+1}^\nu + x_{n+2}^\nu .$$

Bei der Spezialisierung $x_{n+2} = -x_{n+1}$ wird $S_\nu = s_\nu$ (für ungerade ν). Bilden wir daher die Form

$$\varphi(x_1, \ldots, x_n, x_{n+1}, x_{n+2}) = \sum D_{\lambda_1 \lambda_3 \ldots \lambda_{2n+1}} S_1^{\lambda_1} S_3^{\lambda_3} \cdots S_{2n+1}^{\lambda_{2n+1}}$$

mit den obigen Koeffizienten $D_{\lambda_1 \lambda_3 \ldots \lambda_{2n+1}}$, so geht diese bei der Spezialisierung $x_{n+1} + x_{n+2} = 0$ über in das Polynom (I), verschwindet also. Infolgedessen ist sie durch $x_{n+1} + x_{n+2}$ teilbar. Ebenso ist sie aber als symmetrische Form durch jede Summe $x_\nu + x_\mu$, also auch durch das Produkt

$$\prod_{1 \le \nu < \mu \le n+2} (x_\nu + x_\mu)$$

teilbar. Dieses Produkt ist aber vom Grad $(n+1)(n+2)/2 = H$, während die Form φ nur vom kleineren Grad H' ist, also nicht teilbar sein kann, ohne identisch zu verschwinden. Daher ist

$$\sum D_{\lambda_1 \lambda_3 \ldots \lambda_{2n+1}} S_1^{\lambda_1} S_3^{\lambda_3} \cdots S_{2n+1}^{\lambda_{2n+1}} = 0.$$

Das besagt nun, daß zwischen den $n+1$ Potenzsummen $S_1, S_3, \ldots, S_{2n+1}$ von $n+2$ Variablen $x_1, \ldots, x_{n+1}\, x_{n+2}$ eine Abhängigkeit besteht, während sie doch voneinander unabhängig sind (wie oben mit $n-2$ an Stelle von n festgestellt wurde). Die Annahme der Abhängigkeit (I) führt also zu einem Widerspruch.

Um nun die Abhängigkeit der Höhe H zu finden, setzen wir

$$\text{(1)} \qquad F(t) = (1 - x_1 t) \ldots (1 - x_n t) = 1 + a_1 t + a_2 t^2 + \cdots + a_n t^n .$$

Dann ist[2])

$$\frac{F'(t)}{F(t)} = \sum_{\nu=1}^{n} \frac{-x_\nu}{1-x_\nu t} = -\sum_{\nu=1}^{n} (x_\nu + x_\nu^2 t + x_\nu^3 t^2 + \cdots)$$
$$= -(s_1 + s_2 t + s_3 t^2 + \cdots).$$

Wenn man hier t durch $-t$ ersetzt, ergibt sich durch Addition

$$\frac{F'(t)}{F(t)} + \frac{F'(-t)}{F(-t)} = -2(s_1 + s_3 t^2 + s_5 t^4 + \cdots). \tag{2}$$

Multipliziert man mit den Nennern herauf und setzt dann die beiderseits entstehenden Koeffizienten von $t^0, t^2, t^4, \ldots, t^{2n}$ einander gleich, so erhält man $n+1$ in $s_1, s_3, \ldots, s_{2n+1}$ lineare Gleichungen, die aber in bezug auf die a_ν nicht linear sind, so daß sich die a_ν nicht ohne weiteres eliminieren lassen. Zu linearen Gleichungen für die a_ν gelangt man nach einer Idee von PÓLYA auf folgende Weise[3]):

Setzt man

$$\frac{F(-t)}{F(t)} = 1 + w_1 t + w_2 t^2 + w_3 t^3 + \cdots, \tag{3}$$

so ergibt sich durch logarithmische Ableitung mit Rücksicht auf (2):

$$\begin{cases} \dfrac{w_1 + 2w_2 t + 3w_3 t^2 + \cdots}{1 + w_1 t + w_2 t^2 + \cdots} = -\left(\dfrac{F'(t)}{F(t)} + \dfrac{F'(-t)}{F(-t)}\right) \\ \qquad = 2(s_1 + s_3 t^2 + s_5 t^4 + \cdots). \end{cases} \tag{4}$$

Wenn man hier mit dem linken Nenner heraufmultipliziert und dann gleiche Potenzen von t links und rechts vergleicht, erhält man zur Berechnung der w_ν die Rekursionsformeln

$$\begin{cases} w_1 = 2s_1, & 2w_2 = 2s_1 w_1, \\ (2\nu-1)\, w_{2\nu-1} = 2s_1 w_{2\nu-2} + 2s_3 w_{2\nu-4} + \cdots + 2s_{2\nu-3} w_2 + 2s_{2\nu-1}, \\ 2\nu\, w_{2\nu} = 2s_1 w_{2\nu-1} + 2s_3 w_{2\nu-3} + \cdots + 2s_{2\nu-3} w_3 + 2s_{2\nu-1} w_1, \end{cases} \tag{5}$$

wonach klar ist, daß die w_ν Polynome von $s_1, s_3, s_5, \ldots$ sind, und zwar w_ν vom Gewicht ν. Für $\nu \leqq 7$ ist z. B.

$$w_1 = 2s_1, \quad w_2 = 2s_1^2, \quad w_3 = \tfrac{1}{3}(4s_1^3 + 2s_3), \quad w_4 = \tfrac{1}{3}(2s_1^4 + 4s_1 s_3),$$
$$w_5 = \tfrac{1}{15}(4s_1^5 + 20 s_1^2 s_3 + 6 s_5), \quad w_6 = \tfrac{1}{45}(4s_1^6 + 40 s_1^3 s_3 + 10 s_3^2 + 36 s_1 s_5),$$
$$w_7 = \tfrac{1}{315}(8s_1^7 + 140 s_1^4 s_3 + 140 s_1 s_3^2 + 252 s_1^2 s_5 + 90 s_7).$$

Aus (5) ergibt sich weiter durch vollständige Induktion, daß $w_{2\nu-1}$ das Glied $\frac{2s_{2\nu-1}}{2\nu-1}$ enthält und $w_{2\nu}$ für $\nu > 1$ das Glied $\frac{4s_1 s_{2\nu-1}}{2\nu-1}$; ferner enthält w_ν sowohl für gerades als ungerades ν das Glied $\frac{2^\nu s_1^\nu}{\nu!}$, welches lediglich für $\nu = 1$ mit dem zuvor angegebenen zusammenfällt.

[2]) Mit den hier auftretenden Potenzreihen, deren Koeffizienten keine Zahlen sind, so daß Konvergenz nicht in Frage kommt, läßt sich ganz formal rechnen. Man könnte aber (unbequemer) auch mit endlichen Summen und Restglied rechnen, analog wie ich es z. B. in Band 1 meiner „Algebra", S. 160, 2. und 3. Aufl. S. 151f. aus didaktischen Gründen getan habe.

[3]) PÓLYA, G.: Remarques sur un problème d'algèbre étudié par Laguerre. J. de math. pures et appl. **31**, 37—47 (1952).

Nun folgt aus (3) durch Multiplikation mit $F(t)$:

$$(1 + w_1 t + w_2 t^2 + \cdots)(1 + a_1 t + a_2 t^2 + \cdots a_n t^n) = 1 - a_1 t + a_2 t^2 - \cdots + (-1)^n a_n t^n.$$

Vergleicht man hier beiderseits die Koeffizienten von $t^1, t^3, t^5, \ldots, t^{2n+1}$, so kommt

$$(6)\qquad \begin{cases} w_1 + 2a_1 = 0, \\ w_3 + w_2\, a_1 + w_1 a_2 + 2a_3 = 0, \\ w_5 + w_4\, a_1 + w_3 a_2 + w_2 a_3 + w_1 a_4 + 2a_5 = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ w_{2n+1} + w_{2n} a_1 + w_{2n-1} a_2 + \cdots + w_1 a_{2n} + 2a_{2n+1} = 0, \end{cases}$$

wobei $a_\nu = 0$ für $\nu > n$ zu setzen ist. Aus diesen $n+1$ Gleichungen kann man $a_1, a_2, \ldots, a_n$ eliminieren und erhält

$$(7)\qquad \begin{vmatrix} w_1 & 2 & & & & & \\ w_3 & w_2 & w_1 & 2 & & & \\ w_5 & w_4 & w_3 & w_2 & w_1 & 2 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-1} & w_{2n-2} & w_{2n-3} & \cdots & \cdots & \cdots & w_{n-1} \\ w_{2n+1} & w_{2n} & w_{2n-1} & \cdots & \cdots & \cdots & w_{n+1} \end{vmatrix} = 0,$$

wo die leeren Stellen durch Nullen zu besetzen sind. Diese Formel stellt nun, da die w_ν Polynome von $s_1, s_3, s_5, \ldots$ vom Gewicht ν sind, eine Abhängigkeit zwischen den $n+1$ Potenzsummen $s_1, s_3, \ldots s_{2n+1}$ dar, und zwar, wie man am einfachsten aus dem Diagonalglied $w_1 w_2 w_3 \ldots w_{n+1}$ der Determinante (7) feststellt, von der Höhe

$$1 + 2 + 3 + \cdots + (n+1) = \tfrac{1}{2}(n+1)(n+2) = H.$$

Sie ist also die gesuchte Abhängigkeit der Höhe H, vorausgesetzt allerdings, daß die Determinante (7) als Polynom der s_ν nicht etwa identisch verschwindet. Das tut sie aber nicht. Denn sie enthält, da in w_{2n+1} das Glied $\frac{2 s_{2n+1}}{2n+1}$ steckt und die übrigen in der Determinante (7) vorkommenden w_ν von s_{2n+1} unabhängig sind (ihr Index, also ihr Gewicht, ist ja kleiner als $2n+1$), jedenfalls den Bestandteil

$$(8)\qquad \begin{cases} (-1)^n \frac{2 s_{2n+1}}{2n+1} \begin{vmatrix} 2 & & & & & \\ w_2 & w_1 & 2 & & & \\ w_4 & w_3 & w_2 & w_1 & 2 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-2} & w_{2n-3} & \cdots & \cdots & \cdots & w_{n-1} \end{vmatrix} \\ \qquad = (-1)^n \frac{4 s_{2n+1}}{2n+1} \begin{vmatrix} w_1 & 2 & & & & \\ w_3 & w_2 & w_1 & 2 & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-3} & w_{2n-4} & \cdots & \cdots & \cdots & w_{n-1} \end{vmatrix}. \end{cases}$$

Hier enthält nun die rechts stehende Determinante entsprechend den Bestandteil

$$(-1)^{n-2}\frac{4s_{2n-3}}{2n-3}\begin{vmatrix} w_1 & 2 & & \\ w_3 & w_2 & w_1 & 2 \\ \cdots & \cdots & \cdots & \cdots \\ w_{2n-7} & w_{2n-8} & \cdots & w_{n-3} \end{vmatrix}.$$

Durch Fortsetzung dieses Verfahrens erkennt man, daß die Determinante (7) für gerades $n=2r$ das Glied

$$\frac{4s_{4r+1}}{4r+1}\cdot\frac{4s_{4r-3}}{4r-3}\cdots\frac{4s_5}{5}\cdot\frac{2s_1}{1}$$

enthält[4]) und für ungerades $n=2r-1$ das Glied

$$(-1)^r\frac{4s_{4r-1}}{4r-1}\cdot\frac{4s_{4r-5}}{4r-5}\cdots\frac{4s_7}{7}\cdot\frac{4s_3}{3}.$$

Sie verschwindet also als Polynom der s_ν nicht identisch. Unsere Überlegung zeigt zugleich, daß die Determinante (7) den Bestandteil (8) enthält und daß im übrigen s_{2n+1} nicht mehr darin vorkommt. Die Abhängigkeit (7) ist also, wie in Satz 1 behauptet, in bezug auf s_{2n+1} linear.

Es ergibt sich beispielsweise

für $n=1$: $$\begin{vmatrix} w_1 & 2 \\ w_3 & w_2 \end{vmatrix} = 0,$$

oder nach Multiplikation mit $-3/4$: $s_1^3-s_3=0$,

für $n=2$: $$\begin{vmatrix} w_1 & 2 & \\ w_3 & w_2 & w_1 \\ w_5 & w_4 & w_3 \end{vmatrix} = 0,$$

oder nach Multiplikation mit $3\cdot 3\cdot 5/8$:

$$s_1^6-5s_1^3s_3-5s_3^2+9s_1s_5=0,$$

für $n=3$: $$\begin{vmatrix} w_1 & 2 & & \\ w_3 & w_2 & w_1 & 2 \\ w_5 & w_4 & w_3 & w_2 \\ w_7 & w_6 & w_5 & w_4 \end{vmatrix} = 0,$$

oder nach Multiplikation mit $3\cdot 3\cdot 3\cdot 5\cdot 5\cdot 7/16$:

$$s_1^{10}-15s_1^7s_3+63s_1^5s_5-225s_1^3s_7+315s_1^2s_3s_5-175s_1s_3^3+225s_3s_7-189s_5^2=0.$$

[4]) Kein Druckfehler! Im letzten Faktor heißt es $2s_1$, nicht $4s_1$. Denn bei Fortsetzung des Verfahrens stößt man zuletzt auf die einreihige Determinante $|w_1|$, zu der kein Faktor 2 mehr hinzutritt, und es ist $w_1=2s_1$.

§ 2. Ein Satz von Pólya.

Läßt man von den $n+1$ Gleichungen (6) die letzte weg, so kann man die ersten n nach $a_1, a_2, \ldots, a_n$ auflösen. Die dabei auftretende Nennerdeterminante ist, nach den Elementen der ersten Zeile entwickelt, gleich $2\,W_n$, wobei

$$W_n = \begin{vmatrix} w_1 & 2 & & \\ w_3 & w_2 & w_1 & 2 \\ \cdots & \cdots & \cdots & \cdots \\ w_{2n-3} & w_{2n-4} & \cdots & w_{n-1} \end{vmatrix} \qquad (W_1 = 1,\ W_2 = w_1) \tag{9}$$

ist. Daß W_n als Polynom der s_ν nicht identisch verschwindet, kann zunächst genau so gezeigt werden wie oben bei der Determinante (7)[5]. Wir werden es aber gleich noch auf eine zweite Art beweisen, indem wir für W_n als Polynom der x_ν einen anderen Ausdruck herleiten. Wird die Variable x_n zu $-x_{n-1}$ spezialisiert[6]), dann geht s_ν für ungerades ν über in $x_1^\nu + x_2^\nu + \cdots + x_{n-2}^\nu$. Somit geht W_n genau in die Determinante (7) mit $n-2$ an Stelle von n über [man beachte dabei die Fußnote [5])]. Da aber (7) als Polynom der x_ν verschwindet, so besagt das, wenn man die für n gewonnene Erkenntnis auf $n-2$ anwendet, daß W_n bei der Spezialisierung $x_{n-1} + x_n = 0$ verschwindet. Infolgedessen ist W_n durch $x_{n-1} + x_n$ teilbar. Ebenso ist W_n als symmetrische Form auch durch jeden Faktor $x_\nu + x_\mu$ und folglich durch das Produkt

$$P_n = \prod_{1 \le \nu < \mu \le n} (x_\nu + x_\mu) \tag{10}$$

teilbar. Nun ist P_n vom Grad $n(n-1)/2$ und die Determinante W_n ist homogen und, wie man sofort aus dem Diagonalglied $w_1 w_2 \ldots w_{n-1}$ ersieht, ebenfalls vom Grad $n(n-1)/2$. Somit ist

$$W_n = K_n P_n, \tag{11}$$

wo K_n eine Konstante ist, die wir jetzt durch geeignete Spezialisierung der x_ν berechnen wollen.

Im Fall eines ungeraden $n = 2r-1$ spezialisieren wir die x_ν zu den $(2r-1)$-ten Einheitswurzeln, also

$$x_\nu = e^{2\nu\varphi i} \quad (\nu = 1, 2, \ldots, 2r-1), \qquad \text{wobei} \qquad \varphi = \frac{\pi}{2r-1}.$$

Dann ist

$$F(t) = \prod_{\nu=1}^{2r-1} (1 - e^{2\nu\varphi i}\, t) = 1 - t^{2r-1},$$

$$\frac{F(-t)}{F(t)} = \frac{1 + t^{2r-1}}{1 - t^{2r-1}} = \frac{1 + t^n}{1 - t^n} = 1 + 2t^n + 2t^{2n} + \cdots.$$

---

[5]) Man beachte übrigens, daß n auch die Anzahl der Variablen x_ν bezeichnet, so daß die Determinante (7), die aus der Determinante (9) formal dadurch hervorgeht, daß n durch $n+2$ ersetzt wird, keineswegs mit W_{n+2} zu bezeichnen ist. Denn in (7) und (9) sind die w_ν Formen von n Variablen, während sie bei W_{n+2} Formen von $n+2$ Variablen wären.

[6]) Das geht nur für $n > 2$. Aber die Fälle $n = 1$ und $n = 2$ erledigen sich von selbst. Es ist $W_1 = 1$ und $W_2 = w_1 = 2s_1 = 2(x_1 + x_2)$.

Also ist $w_n = 2$, während alle anderen in W_n auftretenden w_ν verschwinden. Die Determinante W_n hat $n-1 = 2r-2$ Reihen und sieht jetzt so aus:

$$\left.\begin{array}{|cccccccc|}
0 & 2 & & & & & & \\
0 & 0 & 0 & 2 & & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & \dots & 0 & 0 & 2 \\
2 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & \dots & 0 & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & \dots & 0 & 2 & 0
\end{array}\right\}\begin{array}{l} r-1 \\ \\ \\ \\ r-1 \\ \\ \\ \\ \end{array}$$

In jeder Zeile und in jeder Spalte steht ein Zweier, und zwar haben die Zweier mit den Zeilennummern $1, 2, \dots, 2r-2$ der Reihe nach die Spaltennummern

$$2, 4, 6, \dots, 2r-2, 1, 3, 5, \dots, 2r-3.$$

Die Anzahl der Inversionen ist

$$1 + 2 + 3 + \cdots + (r-1) = r(r-1)/2,$$

und folglich ist

$$W_{2r-1} = (-1)^{r(r-1)/2} \cdot 2^{2r-2}.$$

Zur Berechnung von P_{2r-1} notieren wir zunächst die aus der Definition von $F(t)$ für $t = -e^{-2\mu\varphi i}$ hervorgehende Formel

$$\prod_{\nu=1}^{2r-1} (1 + e^{2(\nu-\mu)\varphi i}) = 1 + 1 = 2,$$

aus der einerseits nach Multiplikation mit $(e^{2\mu\varphi i})^{2r-1} = 1$, anderseits für $\mu = 0$ und nach Division durch $1 + e^{2(2r-1)\varphi i} = 2$ folgt:

$$\prod_{\nu=1}^{2r-1} (e^{2\mu\varphi i} + e^{2\nu\varphi i}) = 2, \quad \prod_{\nu=1}^{2r-2} (1 + e^{2\nu\varphi i}) = 1. \tag{12}$$

Nunmehr ist

$$P_{2r-1}^2 \cdot \prod_{\nu=1}^{2r-1} (x_\nu + x_\nu) = \prod_{\mu=1}^{2r-1} \prod_{\nu=1}^{2r-1} (x_\mu + x_\nu)$$

$$= \prod_{\mu=1}^{2r-1} \prod_{\nu=1}^{2r-1} (e^{2\mu\varphi i} + e^{2\nu\varphi i}) = \prod_{\mu=1}^{2r-1} 2 = 2^{2r-1},$$

und da das links stehende einfache Produkt gleich

$$\prod_{\nu=1}^{2r-1} (2e^{2\nu\varphi i}) = 2^{2r-1} \cdot e^{2r(2r-1)\varphi i} = 2^{2r-1}$$

ist, ergibt sich zunächst $P_{2r-1}^2 = 1$. Die Berechnung von P_{2r-1} selbst gestaltet sich nun so: Es ist

$$P_{2r-1} = \prod_{\nu<\mu} (e^{2\nu\varphi i} + e^{2\mu\varphi i}) = \prod_{\nu<\mu} e^{2\nu\varphi i} \cdot \prod_{\nu<\mu} (1 + e^{2(\mu-\nu)\varphi i})$$
$$= e^{2a\varphi i} \cdot \prod_{\mu=2}^{2r-1} \prod_{\nu=1}^{\mu-1} (1 + e^{2(\mu-\nu)\varphi i}),$$

wobei es auf die ganze Zahl a nicht ankommen wird, ebenso auch später bei den ganzen Zahlen b, c. Setzt man

$$\prod_{\nu=1}^{\mu-1} (1 + e^{2(\mu-\nu)\varphi i}) = \eta_\mu,$$

so ist also

$$P_{2r-1} = e^{2a\varphi i} \prod_{\mu=2}^{2r-1} \eta_\mu = e^{2a\varphi i} \cdot \eta_{2r-1} (\eta_{2r-2}\eta_2)(\eta_{2r-3}\eta_3) \cdots (\eta_{r+1}\eta_{r-1})\eta_r.$$

Nun ist nach der zweiten Formel (12): $\eta_{2r-1} = 1$; ferner

$$\eta_{r+k}\eta_{r-k} = \prod_{\nu=1}^{r+k-1} (1 + e^{2(r+k-\nu)\varphi i}) \cdot \prod_{\nu=1}^{r-k-1} (1 + e^{2(r-k-\nu)\varphi i})$$
$$= \prod_{\nu=1}^{r+k-1} e^{2(r+k-\nu)\varphi i} \cdot \prod_{\nu=1}^{r+k-1} (e^{2(r-1-k+\nu)\varphi i} + 1) \cdot \prod_{\nu=1}^{r-k-1} (1 + e^{2(r-k-\nu)\varphi i})$$
$$= e^{2b\varphi i} \cdot \prod_{\nu=1}^{2r-2} (1 + e^{2\nu\varphi i}) = e^{2b\varphi i} \quad \text{(nach (12))}.$$

Außerdem ist

$$\eta_r = \prod_{\nu=1}^{r-1} (1 + e^{2(r-\nu)\varphi i}) = \prod_{\nu=1}^{r-1} e^{(r-\nu)\varphi i} \cdot \prod_{\nu=1}^{r-1} (e^{-(r-\nu)\varphi i} + e^{(r-\nu)\varphi i})$$
$$= e^{r(r-1)/2 \cdot \varphi i} \cdot \prod_{\nu=1}^{r-1} \big(2\cos(r-\nu)\varphi\big).$$

Setzt man das alles in den zuletzt gefundenen Ausdruck für P_{2r-1} ein, so ergibt sich

$$P_{2r-1} = e^{2c\varphi i} \cdot e^{r(r-1)/2 \cdot \varphi i} \cdot \prod_{\nu=1}^{r-1} \big(2\cos(r-\nu)\varphi\big).$$

Da die Cosinus offenbar positiv sind und da wir bereits wissen, daß $P_{2r-1}^2 = 1$ ist, kommt also schließlich

$$P_{2r-1} = e^{[2c + r(r-1)/2]\varphi i}.$$

Wegen $P_{2r-1}^2 = 1$ muß $2c + r(r-1)/2$ durch $2r-1$ teilbar sein. Setzt man demgemäß

$$2c + r(r-1)/2 = k(2r-1), \quad \text{so daß} \quad k \equiv r(r-1)/2 \pmod 2$$

ist, so ergibt sich schließlich

$$P_{2r-1} = e^{k(2r-1)\varphi i} = e^{k\pi i} = (-1)^k = (-1)^{r(r-1)/2}.$$

Für die Konstante K_n in (11) findet man also den Wert

$$K_n = K_{2r-1} = \frac{W_{2r-1}}{P_{2r-1}} = 2^{2r-2} = 2^{n-1}.$$

Im Fall eines geraden $n = 2r$ spezialisieren wir $x_{2r} = 0$ und

$$x_\nu = e^{2\nu\varphi i} \quad (\nu = 1, 2, \ldots, 2r-1), \quad \text{wobei} \quad \varphi = \frac{\pi}{2r-1}.$$

Dann ist wieder $F(t) = 1 - t^{2r-1}$ und

$$\frac{F(-t)}{F(t)} = \frac{1 + t^{2r-1}}{1 - t^{2r-1}} = \frac{1 + t^{n-1}}{1 - t^{n-1}} = 1 + 2t^{n-1} + 2t^{2n-2} + \cdots.$$

Daher ist diesmal $w_{n-1} = 2$ und alle anderen in W_n auftretenden w_ν verschwinden. Die Determinante W_n hat $n-1 = 2r-1$ Reihen und sieht so aus:

$$\left.\begin{array}{|cccccccccccc|}
0 & 2 & & & & & & & & & & \\
0 & 0 & 0 & 2 & \ldots & & & & & & & \\
\multicolumn{12}{|c|}{\ldots\ldots\ldots\ldots} \\
0 & 0 & 0 & 0 & \ldots & 0 & 2 & 0 & & & &
\end{array}\right\} r-1$$

$$\left.\begin{array}{|cccccccccccc|}
2 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & & & & \\
0 & 0 & 2 & 0 & \ldots & 0 & 0 & 0 & & & & \\
\multicolumn{12}{|c|}{\ldots\ldots\ldots\ldots} \\
0 & 0 & 0 & 0 & \ldots & 0 & 0 & 2 & & & &
\end{array}\right\} r.$$

Man hat wieder $r(r-1)/2$ Inversionen und folglich ist

$$W_{2r} = (-1)^{r(r-1)/2} \cdot 2^{2r-1}.$$

Für P_n ergibt sich diesmal

$$P_n = P_{2r} = \prod_{\nu=1}^{2r-1} (e^{2\nu\varphi i} + 0) \cdot \prod_{1 \le \nu < \mu \le 2r-1} (e^{2\nu\varphi i} + e^{2\mu\varphi i}).$$

Das erste Produkt ist gleich

$$e^{2r(2r-1)\varphi i} = e^{2r\pi i} = 1,$$

und das zweite ist dasselbe, für das wir oben den Wert $(-1)^{r(r-1)/2}$ fanden. Für die Konstante K_n in (11) ergibt sich also diesmal

$$K_n = K_{2r} = \frac{W_{2r}}{P_{2r}} = 2^{2r-1} = 2^{n-1}.$$

Das Resultat ist daher für gerade und ungerade n das gleiche und es ist $W_n = 2^{n-1} P_n$, was nach Fußnote [6]) auch für den seither ausgeschlossenen Fall $n = 2$ gilt. Somit gewinnen wir die folgende Präzisierung eines Satzes von PÓLYA:

Satz 2. *Sind $s_\nu = x_1^\nu + \cdots + x_n^\nu$ die Potenzsummen von n Variablen, so sind $s_1, s_3, s_5, \ldots, s_{2n-1}$ voneinander unabhängig und jede symmetrische Form ist eine rationale, im allgemeinen gebrochene Funktion von diesen. Insbesondere gilt die Formel*

$$\prod_{\nu<\mu}(x_\nu + x_\mu) = \frac{1}{2^{n-1}} \begin{vmatrix} w_1 & 2 & & & & \\ w_3 & w_2 & w_1 & 2 & & \\ w_5 & w_4 & w_3 & w_2 & w_1 & 2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-3} & w_{2n-4} & w_{2n-5} & \cdots & \cdots & w_{n-1} \end{vmatrix} = \frac{W_n}{2^{n-1}}$$

$$(W_1 = 1,\ W_2 = w_1),$$

während die symmetrischen Grundfunktionen $a_1, a_2, \ldots, a_n$ sich durch Auflösung der ersten n Gleichungen (6) ergeben, wobei die Nennerdeterminante gleich $2\,W_n$ ist. Die w_ν sind dabei Polynome von $s_1, s_3, s_5, \ldots, s_{2n-1}$, die sich rekurrent aus dem Formelsystem (5) berechnen.

Herr Pólya hat nicht die Determinante W_n betrachtet (a. a. O.), sondern diese:

$$R_n = \begin{vmatrix} w_1 & w_2 & \cdots & w_n \\ w_2 & w_3 & \cdots & w_{n+1} \\ & \cdots & \cdots & \\ w_n & w_{n+1} & \cdots & w_{2n-1} \end{vmatrix}.$$

Er konnte zeigen: Erstens, daß R_n als Polynom von $s_1, s_3, \ldots, s_{2n-1}$ in zwei Faktoren zerfällt, deren einer von s_{2n-1} frei und gleich P_n und deren anderer gleich $(-1)^{n(n+1)\cdot 2} \cdot 2^n a_n P_n$ ist, zweitens, daß auch die Produkte $a_\nu P_n$ Polynome von $s_1, s_3, \ldots, s_{2n-1}$ sind; doch hat er für P_n keine formelmäßige Darstellung gegeben (außer für $n = 1, 2, 3, 4$). Vorläufer von Pólya ist Laguerre[7]).

§ 3. Zusätzliche Bemerkungen.

Bemerkung I. W_n ist als Polynom der s_ν irreduzibel (Pólya a. a. O.). Im Fall der Reduzibilität würde nämlich, indem man $n+2$ an Stelle von n schreibt[8]) und dann die Variablen x_{n+1}, x_{n+2} zu 0 spezialisiert, sich auch das Polynom (7) als reduzibel erweisen. Dann müßte einer seiner Faktoren als Polynom der x_ν verschwinden, und das wäre eine Abhängigkeit zwischen $s_1, s_3, \ldots, s_{2n-1}$ von einer Höhe kleiner als H, die es aber nach Satz 1 nicht gibt.

Bemerkung II. Wenn man $a_1, a_2, \ldots, a_n$ durch $s_1, s_2, \ldots, s_n$ ausdrückt, so kommen bei a_ν nur die ersten ν vor: $s_1, \ldots, s_\nu$. Man könnte denken, daß bei der Darstellung durch $s_1, s_3, \ldots, s_{2n-1}$ vielleicht etwas Ähnliches gilt. Das trifft aber nicht zu. Bei a_1 allerdings ist es so, daß nur s_1 vorkommt; es ist ja $a_1 = -s_1$. Bei der Auflösung der ersten n Gleichungen (6) nach a_1 kommt

[7]) Laguerre, E.: Sur un problème d'algèbre. Bull. Soc. Math. France **5**, 26—30 (1877) = Oeuvres **1**, 119—122.

[8]) Das darf man, da für $n = 1$ und $n = 2$ die Sache trivial ist: $W_1 = 1$ ist konstant, $W_2 = w_1 = 2s_1$ ist linear.

eben auch im Zähler die Determinante W_n als Faktor vor, der sich dann gegen den Nenner weghebt. Aber schon bei a_2 ist das nicht so; hier kann kein etwaiger Faktor der Zählerdeterminante gegen einen Faktor des Nenners W_n weggekürzt werden. Denn wegen der Irreduzibilität von W_n müßte W_n selbst dieser Faktor sein, und man erhielte a_2 als Polynom von $s_1, s_3, \ldots, s_{2n-1}$; aber das müßte vom Gewicht 2 sein, könnte also nur die Gestalt $C s_1^2$ haben und wäre nicht gleich a_2.

Bemerkung III. Durch die Potenzsummen $s_1, s_2, \ldots, s_n$ sind die Grundfunktionen eindeutig bestimmt, gleichgültig, ob die x_ν unabhängige Variable sind oder irgendwie spezialisiert. Dabei dürfen für $s_1, s_2, \ldots, s_n$ willkürliche Werte vorgegeben sein, es gibt immer zugehörige $a_1, a_2, \ldots, a_n$, und zwar eindeutig. Durch $s_1, s_3, \ldots, s_{2n-1}$ dagegen sind die a_ν nur dann eindeutig bestimmt, wenn die Determinante W_n von 0 verschieden ist. Wenn dagegen die $s_1, s_3, \ldots, s_{2n-1}$ so vorgegeben sind, daß $W_n = 0$ ist, so gibt es zugehörige Systeme $a_1, \ldots, a_n$ dann und nur dann, und zwar unendlich viele, wenn auch alle Zählerdeterminanten, also schließlich alle Determinanten der Matrix

$$\begin{Vmatrix} w_1 & 2 & & & & & \\ w_3 & w_2 & w_1 & 2 & & & \\ w_5 & w_4 & w_3 & w_2 & w_1 & 2 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{2n-1} & w_{2n-2} & w_{2n-3} & \cdots & \cdots & \cdots & w_{n-1} \end{Vmatrix}$$

verschwinden. Zum Beispiel für $n = 3$ ist

$$W_3 = \begin{vmatrix} w_1 & 2 \\ w_3 & w_2 \end{vmatrix} = \tfrac{1}{3}(s_1^3 - s_3).$$

Wird nun etwa $s_1 = c$, $s_3 = c^3$ vorgegeben, so daß W_3 verschwindet, so ist

$$w_1 = 2c,\ w_2 = 2c^2,\ w_3 = 2c^3,\ w_4 = 2c^4,\ w_5 = 2(4c^5 + s_5)/5,$$

und die obige Matrix ist folgende:

$$\begin{Vmatrix} 2c & & 2 & 0 & 0 \\ 2c^3 & & 2c^2 & 2c & 2 \\ 2(4c^5 + s_5)/5 & 2c^4 & 2c^3 & 2c^2 & \end{Vmatrix}.$$

Ihre Determinanten verschwinden sämtlich nur für $s_5 = c^5$. Gibt man also für s_5 einen von c^5 verschiedenen Wert vor, so gibt es keine zugehörigen a_ν, also kein Polynom $F(t)$; wenn aber $s_5 = c^5$ vorgegeben wird, so gibt es unendlich viele; man findet leicht, daß es die Polynome

$$F(t) = (1 - ct)(1 - dt^2)$$

sind mit beliebigem d.

München 27, Friedrich-Herschel-Str. 11.

(Eingegangen am 13. November 1954.)

Rosenthal, A.
Math. Zeitschr. Bd. 63, S. 31—38 (1955)

On the continuity of functions of several variables.

In memory of Issai Schur.

By
Arthur Rosenthal.

1. Introduction. Let us first consider functions $f(x, y)$ of two variables. It is a well known fact that a function $f(x, y)$ can be continuous along every straight line through the point $p_0 = (x_0, y_0)$ without being continuous at p_0 as a function of (x, y)[1]). Moreover, many years ago, H. Lebesgue[2]) called attention to the fact that $f(x, y)$ can be discontinuous at p_0 even if $f(x, y)$ is continuous along every analytic curve through p_0[3]). Hence one may ask the following questions: (a) By which more general class of curves can the straight lines or the analytic curves be replaced, i.e., on which class of curves through p_0 may continuity of f be assumed without implying the continuity of $f(x, y)$ at p_0 as a function of (x, y)? (b) On which class of curves through p_0 is the continuity of f sufficient to guarantee the continuity of $f(x, y)$ as a function of (x, y)? The discussion of these questions will give the following result:

Theorem 1. *If the single-valued function f is continuous at p_0 along every convex curve through p_0 which is (at least) once differentiable*[4]*, then f is also continuous at p_0 as a function of (x, y). Yet f can be continuous along every curve through p_0 which is (at least) twice differentiable without being continuous at p_0 as a function of (x, y).*

---

[1]) Probably this was first observed by J. Thomae, Abriß einer Theorie der complexen Funktionen, 2nd ed., Halle 1873, p. 15; see Encykl. d. Math. Wiss. II A 1 (A. Pringsheim), footnote 254. The following simple example was given in A. Genocchi-G. Peano, Calcolo differentiale e principii di Calcolo integrale, Torino 1884, No. 123:

$$f(x, y) = \begin{cases} \dfrac{x\,y^2}{x^2 + y^4} & \text{except for } x = y = 0, \\ 0 & \text{for } x = y = 0. \end{cases}$$

[2]) Lebesgue, H.: J. de Math. (6) **1**, 199—200 (1905). Analogous observations concerning maxima and minima were previously made by Scheeffer, L.: Math. Ann. **35**, 545—551 (1890).

[3]) An example of this kind is the following:

$$f(x, y) = \begin{cases} 1 & \text{on the curve } y = e^{-\frac{1}{x^2}} \text{ except for } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

A function $f(x, y)$ of this type which, moreover, is continuous everywhere except for $x = y = 0$ can be defined analogously to our formula (2).

[4]) A straight line segment is considered to be a particular case of a convex curve. Furthermore, a differentiable convex arc is also continuously differentiable [see, e.g., O. Haupt-G. Aumann, Differential- und Integralrechnung, vol. 2, p. 61, Berlin 1938].

The properties of curves, applied here, are to be understood in the following manner: Let the curve C have a tangent at p_0. Take this tangent-line as ξ-axis with p_0 as origin. With respect to this (ξ, η)-coordinate system and in a sufficiently small neighborhood N of p_0, the curve C is to be represented by a single-valued function $\eta = \varphi(\xi)$ which, in the first case [see[4]], has a continuous derivative everywhere in N and, in the second case, has, moreover, a second derivative, at least at p_0.

The proof of Theorem 1 is obtained by means of the following general

Theorem 2. *Let M be a bounded infinite set of points in the plane. Then M contains an infinite subset M^* through which a convex curve C can be laid which is (at least) once differentiable. But there does not always exist an infinite subset of M through which any (at least) twice differentiable curve can be laid.*

In the expression "(at least) once or twice differentiable" the phrase "(at least)" will be omitted in the following.

In § 4 many other applications of this Theorem 2 will be given. In § 5 and 6 the generalizations to functions f of more than two variables will be discussed. All results are quite analogous to the 2-dimensional case; but the differentiable convex arcs (in the plane) will be replaced by differentiable "primitive" arcs (in spaces of more dimensions). Their definition will be given in § 5.

2. Proofs (first part). We now prove the first part of Theorem 2. Let M be a bounded infinite set of points in the plane and let A be a denumerably infinite subset of M with only one point p_0 of accumulation and such that p_0 does not belong to A. We join p_0 with all the points of A by straight rays and consider their limits. Let g be a ray issuing from p_0, which occurs as such a limit. If there are infinitely many points of A on g, then the set of these points can be taken as M^* with g representing C. Otherwise there is a subsequence of points $\{a_\nu\}$ of A, not lying on g, such that the rays $\overrightarrow{p_0 a_\nu}$ converge to g. Then we choose p_0 as the origin of a rectangular coordinate-system with g as the positive x-axis. Let (x_ν, y_ν) be the coordinates of a_ν. By $(a\,b\,x)$ we designate the angle $(\leqq \pi)$ between the directed straight line $\overrightarrow{b\,a}$ and the positive x-axis. Then we can choose a subsequence $\{a_{\nu_k}\}$ of $\{a_\nu\}$ such that $x_{\nu_k} > x_{\nu_{k+1}} > 0$ and such that the angles $(a_{\nu_k}\, p_0 x)$ are acute and the sequence of these angles is strictly decreasing. The angles $(a_{\nu_k} a_{\nu_{k+m}} x)$ (for $m = 1, 2, \ldots$) are greater than the angle $(a_{\nu_k} p_0 x)$, but converge to it as $m \to \infty$. Hence from $\{a_{\nu_k}\}$ we can select a new subsequence, which we call $\{b_n\}$, such that the angles $(b_n b_{n+1} x)$ strictly decrease to 0 as n increases. Then the points b_n are the vertices of a generalized convex polygon with denumerably many sides converging to p_0 and with the positive x-axis as half-tangent at p_0. At every vertex b_n (with $n \geqq 2$) we draw the bisector g_n of the supplement of the interior angle $b_{n-1} b_n b_{n+1}$. Let the intersection of g_n and g_{n+1} be designated by c_n. The acute angle γ_n determined by g_n and the x-axis is the arithmetic mean of the angles $(b_{n-1} b_n x)$ and $(b_n b_{n+1} x)$, and hence the γ_n

decrease strictly to 0 as $n \to \infty$. Therefore the generalized polygon $\mathfrak{P}$ with the vertices c_n $(n = 2, 3, \ldots)$ is also convex. Its sides $c_n c_{n+1}$ (containing δ_{n+1}) converge to p_0, again with the positive x-axis as half-tangent of $\mathfrak{P}$ at p_0. If we round off every vertex c_n of $\mathfrak{P}$ by a sufficiently small circular arc which touches the sides of the vertex c_n, then we obtain a differentiable convex arc C with p_0 as endpoint and the positive x-axis as half-tangent at p_0. Joining C with its reflection in the y-axis gives a differentiable convex arc with the x-axis as tangent at p_0. This proves the first part of theorem 2.

The first part of theorem 1 now follows immediately. In fact, let the values of f be continuous at $p_0 \equiv (x_0, y_0)$ on every differentiable convex curve through p_0. Then if $f(x, y)$ were not continuous at p_0 as a function of (x, y), there would be a sequence of points $\{a_\nu\}$ converging to p_0 for which $\lim_{\nu \to \infty} f(x_\nu, y_\nu)$ would exist, but would be different from $f(x_0, y_0)$. Then, according to the above discussion, there would be a subsequence $\{a_{\nu_n}\}$ through which one could lay a differentiable convex curve C. Hence, by assumption, we would have $\lim_{n \to \infty} f(x_{\nu_n}, y_{\nu_n}) = f(x_0, y_0)$, in contradiction to the selection of the $\{a_\nu\}$.

3. Proofs (second part). We shall now show that a bounded infinite set M does not always contain an infinite subset through which a twice differentiable curve can be laid. For this purpose we consider a continuous curve $y = f(x)$ passing through 0 with $f'(0) = 0$ and with the improper second derivative $f''(0) = +\infty$; e.g., for $y = f(x)$ we take the semicubical parabola $y = a \cdot |x|^{\frac{3}{2}}$ (with $a > 0$). On the branch of this curve for which $x > 0$ let $\{a_\nu\}$ be a sequence of points converging to the origin and let this sequence be our set M. Then no twice differentiable curve C can be laid through any infinite subset of M. For let $y = \varphi(x)$ with $\varphi(0) = 0$ and $\varphi'(0) = 0$ represent such a curve C. Then since $\lim_{h \to 0^+} \frac{\varphi'(h) - \varphi'(0)}{h} = L$ (finite), for every $\varepsilon > 0$ there exists a $\delta_1 > 0$ such that for $0 < h < \delta_1$ we have $|L| - \varepsilon < \left| \frac{\varphi'(h)}{h} \right| < |L| + \varepsilon$. On the other hand, to $Q > |L| + \varepsilon$ there exists a $\delta_2 > 0$ such that for $0 < h < \delta_2$ we have $\frac{f'(h)}{h} > Q$ (i.e., in our case, $\frac{3}{2} a \cdot h^{-\frac{1}{2}} > Q$). Thus if $\delta = \min(\delta_1, \delta_2)$, for $0 < h < \delta$ we have $|\varphi'(h)| < f'(h)$, and hence by integration $|\varphi(h)| < f(h)$ for $0 < h < \delta$ (where if f is given, δ depends on φ). Therefore no twice differentiable curve $y = \varphi(x)$ can go through infinitely many points of M. This proves the second part of theorem 2.

The second part of theorem 1 is then an immediate consequence, since one can now easily give examples of functions $f(x, y)$ which are continuous along every twice differentiable curve through $(0, 0)$ without being continuous at the origin as functions of (x, y). For instance set

$$(1) \qquad f(x, y) = \begin{cases} 1 & \text{on the curve } y = |x|^{\frac{3}{2}} \text{ except for } x = 0, \\ 0 & \text{otherwise } [(0, 0) \text{ included}]. \end{cases}$$

In order to obtain such a function which is continuous everywhere except for $(0, 0)$, one can proceed in the following manner. Let C_1, C_2, C_3 be three

curves with the equations $y=a|x|^{\frac{3}{2}}$, $y=b|x|^{\frac{3}{2}}$, $y=c|x|^{\frac{3}{2}}$, respectively, where $b>c>a>0$. In the "strip" S bounded by C_1 and C_2, let there be a "mountain" above the xy-plane whose crest lies above C_3 with a constant altitude h, except at the origin, and which slopes down to the curves C_1 and C_2. We now define the single-valued function $z=f(x, y)$ correspondingly: Inside S it is represented by the "mountain", while it vanishes on C_1 and C_2 (the origin included) and everywhere outside of S. A very simple definite example of such a function $f(x, y)$ is the following:

$$(2)\quad f(x,y)=\begin{cases} |(y|x|^{-\frac{3}{2}}-a)\,(y|x|^{-\frac{3}{2}}-b)|-(y|x|^{-\frac{3}{2}}-a)\,(y|x|^{-\frac{3}{2}}-b) & \text{for } x\neq 0,\\ 0 & \text{for } x=0 \end{cases}$$

with $b>a>0$. Here the curve C_3 is determined by $y=\frac{1}{2}(a+b)|x|^{\frac{3}{2}}$ and on C_3, except for $x=0$, f has the constant value $h=\frac{1}{2}(b-a)^2$.

These examples accomplish what we desire. In fact, every twice differentiable curve C passing through the origin 0 lies outside of S in a sufficiently small deleted neighborhood N of 0, and hence f vanishes everywhere on C in N. This is also the case for every such curve C whose tangent at 0 is the x-axis, since by the above such a C (except for 0) lies below C_1 in a sufficiently small neighborhood of 0.

We also note that a function $F(x, y)$ may be discontinuous at the points of a set dense in a region R, although F is continuous on every curve K in R which is twice differentiable at each of its points. Example: Let (x_n, y_n) $(n=1, 2, \ldots)$ be a sequence of all the points in R with both coordinates rational and set

$$(3)\qquad F(x,y)=\sum_{n=1}^{\infty}\frac{1}{2^n}f(x-x_n, y-y_n)$$

where $f(x, y)$ denotes the function defined in (2).

4. Other applications. Theorem 2 allows many other applications also.

(a) Let $z=F(x, y)$ represent a continuous surface S and let p_0 be a point on S. It is possible that every curve on S through p_0, such that the projection on the xy-plane is a twice differentiable curve, has a tangent at p_0 and that all these tangents lie in a plane, but that, nevertheless, this plane is not a *tangent plane* of S at p_0. Yet if the same holds for every curve on S through p_0 which is projected on the xy-plane as a differentiable convex curve, then that plane is a tangent plane of S at p_0.

An example for the first part of this statement is given by

$$(4)\qquad F(x,y)=f(x,y)\cdot|x|,$$

where $f(x, y)$ is the function defined by (2).

(b) A continuous function $w=\varphi(z)$, where $z=x+iy$, can have a *derivative* with the same value at z_0 along every twice differentiable curve through z_0, although $\varphi'(z_0)$ does not exist. But $\varphi'(z_0)$ certainly exists if $\varphi(z)$ has a derivative with the same value at z_0 along every differentiable convex curve.

An example for the first part of this statement is obtained from the function $F(x, y)$, defined in (4), by setting $\varphi(x+iy)=F(x, y)$.

(c) If the single-valued function $g(x, y)$ is bounded along every differentiable convex curve through the point p_0, then g is *bounded* in a sufficiently small neighborhood of p_0. Yet g can be bounded along every twice differentiable curve through p_0 without being bounded in any neighborhood of p_0.

An example for the last part of this statement is given by

$$g(x, y)=\begin{cases} f(x, y)\cdot\frac{1}{|x|} & \text{for } x\neq 0, \\ 0 & \text{for } x=0, \end{cases} \tag{5}$$

where $f(x, y)$ again designates the function defined by (2).

(d) Let the single-valued function $f(x, y)$ be defined in a deleted neighborhood of $p_0\equiv(x_0, y_0)$. Then

$$\lim_{(x,y)\to(x_0y_0)} f(x, y) \tag{6}$$

certainly exists if f has the *same* limit at p_0 along every differentiable convex curve through p_0. But even if f has the same limit at p_0 along every twice differentiable curve through p_0, then the limit (6) need not exist.

Here it is essential that f has always the *same* limit at p_0. For even the existence of a limit of f at p_0 (but not necessarily always the same limit) along each differentiable curve through p_0 does not guarantee the existence of (6). Example: Let S be the part of the screw surface

$$x=r\cos\vartheta, \qquad y=r\sin\vartheta, \qquad z=\vartheta, \qquad 0\leq\vartheta\leq\frac{\pi}{2}$$

and let $z=f(x, y)$ represent the surface obtained by continuing S by its reflections in the xz- and yz-planes.

(e) The single-valued function $f(x, y)$ has a *maximum* at p_0 if f has a maximum at p_0 along every differentiable convex curve through p_0. But f can have a maximum at p_0 along every twice differentiable curve through p_0 without having a maximum at p_0 as a function of (x, y). Example:

$$f(x, y)=(y-a|x|^{\frac{3}{2}})\cdot(b|x|^{\frac{3}{2}}-y) \quad \text{with } b>a>0. \tag{7}$$

(f) Let $\{f_\nu(x, y)\}$ be a sequence of single-valued functions defined in a neighborhood of p_0. Then the *uniform convergence* of $\{f_\nu(x, y)\}$ at p_0 is implied by the uniform convergence of $\{f_\nu\}$ at p_0 along every differentiable convex curve through p_0; but it is not necessarily implied by the uniform convergence of $\{f_\nu\}$ at p_0 along every twice differentiable curve through p_0.

Proof of the first part: Suppose $f_\nu(x, y)$ is not uniformly convergent at p_0. Then there exists an $\varepsilon_0>0$, a sequence of points $p_n\equiv(x_n, y_n)$ converging to p_0, and, for every n, two indices ν_n and ν_n^* with $\nu_n^*>\nu_n\geq n$, such that

$$|f_{\nu_n^*}(x_n, y_n)-f_{\nu_n}(x_n, y_n)|\geq\varepsilon_0.$$

But then there is a differentiable convex arc C through a subsequence of $\{p_n\}$ along which $\{f_\nu\}$ is not uniformly convergent at p_0.

Example for the second part: Set

$$\psi_\nu(\varrho) = \begin{cases} 2\nu\varrho & \text{for } 0 \leq \varrho \leq \frac{1}{2\nu}, \\ 2(1-\nu\varrho) & \text{for } \frac{1}{2\nu} \leq \varrho \leq \frac{1}{\nu}, \\ 0 & \text{for } \varrho \geq \frac{1}{\nu} \end{cases} \tag{8}$$

and

$$f_\nu(x, y) = f(x, y) \cdot \psi_\nu\left(\sqrt{x^2+y^2}\right) \qquad (\nu = 1, 2, \ldots) \tag{9}$$

where $f(x, y)$ is the function defined by (2).

Moreover, we give an example of a sequence $\{F_\nu(x, y)\}$ of continuous functions in a region R, uniformly convergent along every curve in R which is twice differentiable at each of its points, without being uniformly convergent in any subregion of R: Let (x_n, y_n) $(n = 1, 2, \ldots)$ be a sequence of all the points with both coordinates rational. Then we set

$$F_\nu(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot f_\nu(x - x_n, y - y_n) \qquad (\nu = 1, 2, \ldots) \tag{10}$$

where the $f_\nu(x, y)$ are the functions defined by (9).

(g) Let $\{f_\nu(x, y)\}$ be a sequence of single-valued functions defined in a bounded region R. If $\{f_\nu\}$ is uniformly bounded along every differentiable convex curve in R, then $\{f_\nu\}$ is *uniformly bounded* in R. But $\{f_\nu\}$ can be uniformly bounded along every curve in R which is twice differentiable at each of its points without being uniformly bounded in R.

Example for the second part:

$$f_\nu(x, y) = \begin{cases} f(x, y) \cdot (x^2+y^2)^{-\frac{1}{4}} \cdot \psi_\nu\left(\sqrt{x^2+y^2}\right) & \text{except for } x = y = 0, \\ & (\nu = 1, 2, \ldots) \\ 0 & \text{for } x = y = 0 \end{cases} \tag{11}$$

where $f(x, y)$ and $\psi_\nu(\varrho)$ are the functions defined by (2) and (8).

(h) The situation is quite different for a sequence $\{f_\nu(z)\}$ of regular *analytic* functions in a region R. In order that $\{f_\nu(z)\}$ be *uniformly bounded* in every bounded, closed subregion of R, it is necessary and sufficient that $\{f_\nu(z)\}$ be uniformly bounded *on every closed straight segment of* R. For then $\{f_\nu(z)\}$ is uniformly bounded on every simple closed polygon $\mathfrak{P}$ contained in R, and hence, by the maximum modulus theorem, also in the closed region bounded by $\mathfrak{P}$.

The analogous statement holds also for the *uniform convergence* of $\{f_\nu(z)\}$. This results either from the preceding by means of the Vitali theorem or, quite directly, using the uniform convergence of $\{f_\nu(z)\}$ on $\mathfrak{P}$ and then concluding[5]), by means of the Cauchy integral formula, that $\{f_\nu(z)\}$ converges uniformly also in the region bounded by $\mathfrak{P}$.

[5]) Cf. RUNGE, C.: Acta math. **6**, 232—234 (1884). — BIEBERBACH, L.: Lehrbuch der Funktionentheorie, vol. I, p. 166—167, Leipzig u. Berlin 1921 (4th. ed., 1934, p. 168—169).

5. Primitive arcs. Here and in § 6 we shall discuss the generalization to functions of more than two variables. For this purpose we first define (by induction for $n \geqq 2$) a *primitive arc*, which in R_n will play the same role as the convex arc in the plane. In R_2 a primitive arc is a convex arc. Suppose the notion of a primitive arc in R_{n-1} has been defined; then we define it in R_n in the following manner: Let C be an arc in R_n through the point p_0 with the tangent t_0 at p_0. Let $\pi_0^{(n-1)}$ be the R_{n-1} through p_0 perpendicular to t_0. Then the orthogonal projection C' of C on $\pi_0^{(n-1)}$ is assumed to be a primitive arc with a tangent t_1 at p_0, and the orthogonal projection C'' of C on the plane (t_0, t_1) [6]) is assumed to be a convex arc. The special case is admitted in which C' shrinks to the point p_0; in this case C is a straight line segment (of t_0). According to this recursive definition there are $n-1$ orthogonal planes through p_0, on which the projections of C are convex arcs (provided that none of them shrinks to p_0).

The primitive arc C is *differentiable* if C has a tangent at each of its points. This is the case for a sufficiently small C if and only if each of the projections on the orthogonal planes referred to above is a differentiable convex arc. In fact, C is then even continuously differentiable (i.e., the tangent varies continuously).

We now state what we mean by a *twice differentiable* arc C in R_n. Let C be a continuously differentiable arc through p_0 with the tangent t_0 at p_0. In a sufficiently small neighborhood N of p_0, the projection C' of C on any plane π through t_0 is a continuously differentiable arc. Taking t_0 as ξ-axis with origin at p_0, let a (ξ, η)-coordinate system be introduced in π; then C' is represented in N by a single-valued, continuously differentiable function $\eta = \varphi(\xi)$. We call C "twice differentiable" if, in each such plane π, at least $\varphi''(0)$ exists.

6. Generalization to n dimensions. Now we generalize Theorem 2 for the n-dimensional space R_n.

Theorem 3. *Let M be a bounded infinite set of points in R_n. Then M contains an infinite subset M^* through which a differentiable primitive arc C can be laid. But there does not always exist an infinite subset of M through which any twice differentiable arc can be laid.*

Proof: Let A be a denumerable subset of M with only one point p_0 of accumulation and let p_0 not belong to A. As at the beginning of § 2 either there is a line g containing infinitely many points of A (then g can be taken as C) or otherwise we can select a subsequence $\{a_\nu^1\}$ of A such that the rays $\overline{p_0 a_\nu^1}$ converge to a limiting ray g_1. Then consider the R_{n-1} perpendicular to g_1 at p_0 and the orthogonal projections $\tilde{a}_\nu^1$ of the a_ν^1 on this R_{n-1}. There is a subsequence $\{a_\nu^2\}$ of $\{\tilde{a}_\nu^1\}$ such that the rays $\overline{p_0 a_\nu^2}$ converge to a limiting ray g_2. (The degenerate case that infinitely many $\tilde{a}_\nu^1$ lie on a ray will be discussed subsequently.) Then consider the R_{n-2} (contained in our R_{n-1}) perpendicular

[6]) This is a "tangential osculating plane" of C at p_0; [cf. O. Haupt, J. f. Math. **170**, 157 (1934)].

to g_2 at p_0 and the orthogonal projections $\tilde{a}_\nu^2$ of the a_ν^2 on this R_{n-2}. Again there is a subsequence $\{a_\nu^3\}$ of the $\{\tilde{a}_\nu^2\}$ such that the rays $\overrightarrow{p_0 a_\nu^3}$ converge to a limiting ray g_3. In this way we continue and finally come to a sequence $\{a_\nu^{n-1}\}$ of projections on a plane R_2 perpendicular to g_{n-2} at p_0, such that the rays $\overrightarrow{p_0 a_\nu^{n-1}}$ converge to a limiting ray g_{n-1}. Then let g_n be a ray in this R_2, perpendicular to g_{n-1} at p_0. The subsequence of $\{a_\nu^1\}$ which is projected into $\{a_\nu^{n-1}\}$ is designated by $\{q_\nu\}$.

It is possible that the preceding procedure stops at an earlier moment. There may be a first ray $g_k (2 \leqq k \leqq n-1)$ which contains infinitely many of the projections $\tilde{a}_\nu^{k-1}$; then we take these points as the subsequence $\{a_\nu^k\}$. The corresponding subsequence of $\{a_\nu^1\}$ (which we may call $\{q_\nu\}$ again) is then already contained in the k-dimensional subspace of R_n which is spanned by $g_1, g_2, \ldots, g_k$. Hence our procedure stops here, and we have to consider only the k-dimensional case.

If our procedure does not stop at an earlier stage, then the rays $g_1, g_2, \ldots, g_n$ determine a rectangular $(x_1, x_2, \ldots, x_n)$-coordinate system in R_n. The projections of $\{q_\nu\}$ on the (x_k, x_{k+1})-plane are denoted by $\{q_\nu^k\}$, $(k = 1, 2, \ldots, n-1)$. [Hence $\{q_\nu^{n-1}\}$ is the same as $\{a_\nu^{n-1}\}$.] According to theorem 2, in the (x_1, x_2)-plane we lay a differentiable convex curve C_1 through an infinite subsequence $\{b_\nu^1\}$ of $\{q_\nu^1\}$. Let $\{q_\nu^{2*}\}$ be the subsequence of $\{q_\nu^2\}$, corresponding to $\{b_\nu^1\}$. Then in the (x_2, x_3)-plane we lay a differentiable convex curve C_2 through an infinite subsequence $\{b_\nu^2\}$ of $\{q_\nu^{2*}\}$. Let $\{q_\nu^{3*}\}$ be the subsequence of $\{q_\nu^3\}$, corresponding to $\{b_\nu^2\}$. In this manner we continue until in the (x_{n-1}, x_n)-plane we obtain a differentiable convex curve C_{n-1} through an infinite subsequence $\{b_\nu^{n-1}\}$ of a subsequence $\{q_\nu^{n-1*}\}$ of $\{q_\nu^{n-1}\}$. Let $\{p_\nu\}$ be an infinite subsequence of $\{q_\nu\}$ whose projections on each (x_k, x_{k+1})-plane are included in the $\{b_\nu^k\}$ $(k = 1, 2, \ldots, n-1)$. Then by $C_1, C_2, \ldots, C_{n-1}$ a differentiable primitive arc C is determined going through the points $\{p_\nu\}$. In fact, C is the intersection of $n-1$ projecting cylinders.

The second part of theorem 3 is shown by the following example: Let C be the curve in R_n, which is determined by the curves $x_{k+1} = |x_k|^{\frac{3}{2}}$ in the (x_k, x_{k+1})-planes $(k = 1, 2, \ldots, n-1)$. Then let M be a sequence of points on C, which converges to the origin.

An immediate consequence of Theorem 3 is the following

Theorem 4. *One can generalize Theorem 1 to single-valued functions $f(x_1, x_2, \ldots, x_n)$ if there one replaces "convex curve" by "primitive curve".*

One also obtains easily other applications of Theorem 3, analogous to those discussed in § 4.

Dept. of Math., Purdue University, Lafayette, Indiana (U.S.A.).

(Eingegangen am 12. Dezember 1954.)

KNOPP, K.
Math. Zeitschr. Bd. 63, S. 39—52 (1955)

Nörlund-Verfahren für Funktionen*).

(Zweite Mitteilung.)

Dem Andenken an ISSAI SCHUR gewidmet.

Von

KONRAD KNOPP.

1. Einleitung.

Die funktionalen Analoga zu den bekannten NÖRLUNDschen Verfahren zur Limitierung von Zahlenfolgen sind in einer vorangehenden gemeinsamen Arbeit des Verfassers mit B. VANDERBURG [*I*][1]) nach verschiedenen Richtungen untersucht worden. Dort wurde die Klasse S aller (komplexwertigen) Funktionen $s(t)$ eingeführt, die in $t>0$ erklärt und in jedem Intervall $(0, c)$, $c>0$, meßbar und beschränkt sind, sowie die Klasse $\mathcal{P}$ aller (reellwertigen) Gewichtsfunktionen $p(t)$, die in $t>0$ erklärt, in jedem Intervall (a, b), $0<a<b<\infty$, meßbar und beschränkt sind und für die die Integrale

$$P(x)=\int_0^x p(t)\,dt \tag{1.1}$$

bei 0 absolut konvergieren und für alle $x>0$ positiv sind. Wenn $s\in S$ und $p\in\mathcal{P}$, so heißt die für alle $x>0$ erklärte Funktion

$$\sigma(x)=\frac{1}{P(x)}\int_0^x p(x-t)\,s(t)\,dt \tag{1.2}$$

das N_p-*Mittel* von $s(t)$. Strebt $\sigma(x)\to s$ für $x\to\infty$, so heißt $s(t)$ N_p-*limitierbar* zum Werte s für $t\to\infty$. Dies Verfahren heißt *konvergenztreu*, wenn die Konvergenz von $s(t)$ diejenige von $\sigma(x)$ zur Folge hat[2]), es heißt *multiplikativ*, wenn überdies $\lim\sigma(x)=a\cdot\lim s(t)$ ist, und *regulär*, wenn der Faktor $a=1$ ist. Ist $p(t)\geqq 0$, so sagen wir kurz, N_p sei positiv.

Es werde ferner die Teilklasse $\mathcal{V}<S$ aller Funktionen $s(t)\in S$ eingeführt, die in $(0,\infty)$ von beschränkter Gesamtschwankung sind. Wenn (1.2) jedes $s(t)\in\mathcal{V}$ in ein $\sigma(x)\in\mathcal{V}$ transformiert, so werden die eben erklärten Verfahren *absolut* konvergenztreu, multiplikativ bzw. regulär genannt.

---

*) This research was supported by the United States Air Force through the Office of Scientific Research, of the Air Research and Development Command, under Contract Number AF 18(600)—691.

[1]) Auf diese Arbeit wird im folgenden mit „I" verwiesen.

[2]) Die Grenzwerte sind, wenn nichts anderes ausdrücklich gesagt wird, stets für $t\to\infty$ bzw. $x\to\infty$ genommen.

Liegt ein Integral $\int_0^\infty a(u)\,du$ vor mit $a \in S$ und ist $\int_0^t a(u)\,du = s(t)$ N_p-limitierbar zum Werte s, so heißt das vorgelegte Integral *N_p-summierbar* zum Werte s, insbesondere *absolut* N_p- oder $|N_p|$-summierbar zum Werte s, wenn noch $\sigma(x) \in \mathcal{V}$ gilt.

In § 2 werden zwei Sätze über allgemeine Integraltransformationen bewiesen und in § 3 zwei Sätze über absolute Inklusion (Erklärung s. § 3), die selbständiges Interesse haben. Sie setzen uns instand, in § 4 mehrere Multiplikationssätze zu beweisen, d.h. Sätze, die unter verschiedenen Voraussetzungen auf folgende Frage antworten: Wenn die Integrale $\int_0^\infty a(t)\,dt$ und $\int_0^\infty b(t)\,dt$ je durch ein bestimmtes unserer Verfahren summierbar sind und wenn

$$a * b = \int_0^t a(t-u)\,b(u)\,du = c(t) \tag{1.3}$$

gesetzt wird[3]), — was läßt sich dann über die Summierbarkeit von $\int_0^\infty c(t)\,dt$ sagen? Es werden drei Antworten auf diese Frage gegeben.

Für den Fall unendlicher Reihen $\sum a_\nu$, $\sum b_\nu$, $\sum c_\nu$, bei denen $c_\nu = \sum_{\varrho=0}^{\nu} a_{\nu-\varrho} b_\varrho$ ist, wurden die Antworten von FL. M. MEARS [2], [3], [4], [5] gegeben. Die folgenden Beweise lehnen sich an eine vereinfachte, noch unveröffentlichte Darstellung von G. G. LORENTZ und dem Verfasser an.

2. Zwei Sätze über Integraltransformationen.

Wir brauchen zwei Sätze über Integraltransformationen, die auch selbständiges Interesse haben.

Satz 1. *Die Integraltransformation*[4])

$$\sigma(x) = \int_0^\infty a(x,t)\,s(t)\,dt, \qquad s(t) \in \mathcal{V}, \tag{2.1}$$

ist absolut konvergenztreu, bildet also $\mathcal{V}$ in $\mathcal{V}$ ab, wenn die beiden folgenden Bedingungen erfüllt sind:

$$\int_0^\infty a(x,t)\,dt = A(x) \quad \textit{existiert für jedes} \quad x \geq 0, \tag{2.2}$$

und wird

$$\int_t^\infty a(x,u)\,du = \varphi(x,t), \qquad (x,t \geq 0), \tag{2.3}$$

[3]) Vgl. I, § 3.

[4]) Das Integral rechts in (2.1) und in analogen Fällen ist als $\lim \int_0^T$ für $T \to \infty$ zu verstehen.

gesetzt, so ist die Gesamtschwankung von $\varphi(x,t)$ über $(0,\infty)$ bezüglich x gleichmäßig beschränkt in t, d.h. es gibt ein festes $M>0$, so daß

$$\int_0^\infty |d_x\varphi(x,t)| \leqq M \quad \text{für alle} \quad t \geqq 0 \text{ [5]}. \tag{2.4}$$

Beweis. Da $s(t) \in \mathcal{V}$, ergibt sich durch partielle Integration

$$\sigma(x) = \int_0^\infty \varphi(x,t)\, ds(t) + s(0)\, A(x). \tag{2.5}$$

Auf Grund des Lemmas in I, § 6 und weil $A(x) = \varphi(x,0)$ ist, ergibt sich nun sofort

$$\int_0^\infty |d\sigma(x)| \leqq M\left\{|s(0)| + \int_0^\infty |ds(t)|\right\}. \tag{2.6}$$

Daher ist $\sigma(x) \in \mathcal{V}$.

Bemerkungen. Aus den Voraussetzungen von Satz 1 folgt:

1. Es existiert

$$\lim A(x) = a. \tag{2.7}$$

Denn die Gesamtschwankung von $A(x) = \varphi(x,0)$ über $(0,\infty)$ ist $\leqq M$.

2. Für jedes feste $t \geqq 0$ existiert

$$\lim_{x\to\infty} \int_0^t a(x,u)\, du = a^*(t) \text{ [6]}. \tag{2.8}$$

Denn da die Gesamtschwankung von $\varphi(x,t)$ bezüglich x endlich ist, existiert zunächst $\lim \varphi(x,t)$ für $x\to\infty$ bei jedem festen $t \geqq 0$. Daraus folgt aber die Behauptung, weil die linke Seite von (2.8) mit $\lim\big(\varphi(x,0) - \varphi(x,t)\big)$ für $x\to\infty$ äquivalent ist.

3. Wenn überdies für jedes $T>0$ der Grenzwert (2.8) in $0 \leqq t \leqq T$ gleichmäßig existiert und wenn $|\varphi(0,t)| \leqq M$ [7]) ist für alle $t \geqq 0$, dann ist

$$\lim_{x\to\infty} \sigma(x) = a \cdot s - \int_0^\infty a^*(t)\, ds(t). \tag{2.9}$$

Nach (2.4) ist nämlich für jedes $X>0$ und alle $t \geqq 0$

$$\int_0^X |d_x\varphi(x,t)| \leqq M,$$

folglich

$$|\varphi(X,t) - \varphi(0,t)| \leqq M,$$

---

[5]) Die beiden Bedingungen (2.2) und (2.4) sind auch notwendig dafür, daß (2.1) absolut konvergenztreu ist. Einen Beweis dafür hat K. ZELLER für den Verfasser ausgearbeitet. Doch da im folgenden von der Notwendigkeit der Bedingungen kein Gebrauch gemacht wird, unterdrücken wir den nicht ganz einfachen Beweis. — Vgl. hierzu auch: G. SUNOUCHI und T. TSUCHIKURA [6].

[6]) Im Fall von Matrixtransformationen entspricht dies der Existenz der Spaltenlimites.

[7]) O. B. d. A. darf hier das gleiche M wie in (2.4) genommen werden.

also

$$|\varphi(x,t)| \leq 2M \quad \text{für alle} \quad x,t \geq 0$$

und

$$\left| \int_0^t a(x,u)\,du \right| = |\varphi(x,0) - \varphi(x,t)| \leq 4M.$$

Daher sind alle Voraussetzungen des nachfolgenden Satzes 2 erfüllt [mit $4M$ statt M in (2.10)], so daß die jetzige Behauptung (2.9) in der Behauptung (2.13) dieses Satzes 2 enthalten ist, den wir sogleich beweisen.

4. Wenn $a^*(t) \equiv 0$ ist, so ist (2.1) absolut multiplikativ mit dem Faktor $a = \lim A(x)$ und absolut regulär, wenn noch $a = 1$ ist.

Satz 2. *Die Funktion $\sigma(x)$ in (2.1) konvergiert für jedes $s(t) \in \mathcal{V}$, falls mit einem von x und t unabhängigen M*

$$\left| \int_0^t a(x,u)\,du \right| \leq M \tag{2.10}$$

ist für alle $x,t \geq 0$,

$$A(x) = \int_0^\infty a(x,t)\,dt$$

für alle $x \geq 0$ existiert und

$$A(x) \to a \quad \textit{strebt für} \quad x \to \infty \tag{2.11}$$

und wenn für jedes $T > 0$ der Grenzwert

$$\lim_{x\to\infty} \int_0^t a(x,u)\,du = a^*(t) \tag{2.12}$$

gleichmäßig in $0 \leq t \leq T$ existiert. Sind diese Bedingungen erfüllt, so ist überdies

$$\lim_{x\to\infty} \sigma(x) = a \cdot s - \int_0^\infty a^*(t)\,ds(t) = \sigma. \tag{2.13}$$

Beweis. Mit der Bezeichnung (2.3) ist zunächst nach (2.10) und (2.11)

$$|\varphi(x,t)| \leq 2M \quad \text{für alle} \quad x,t \geq 0$$

und es strebt für $x \to \infty$

$$\varphi(x,t) \to a - a^*(t) \tag{2.14}$$

gleichmäßig in $0 \leq t \leq T$, so daß auch

$$|a - a^*(t)| \leq 2M \quad \text{ist für alle} \quad t > 0. \tag{2.15}$$

Wird also bei gegebenem $\varepsilon > 0$ das T so gewählt, daß

$$\int_T^\infty |ds(t)| < \frac{\varepsilon}{8M}$$

ist und wird

$$\begin{aligned}\sigma(x) - \sigma &= \int_0^T + \int_T^\infty \{\varphi(x,t) - a + a^*(t)\}\,ds(t) + \{A(x) - a\}\,s(0)\\ &= A + B + C\end{aligned}$$

gesetzt, so ist zunächst $|B|<\varepsilon/2$. Ferner ist

$$A=\int_0^T\left\{A(x)-\int_0^t a(x,u)\,du-a+a^*(t)\right\}ds(t),$$

also

$$|A|\leqq|A(x)-a|\int_0^\infty|ds(t)|+\int_0^T\left|\int_0^t a(x,u)\,du-a^*(t)\right|\cdot|ds(t)|<\varepsilon/4$$

für alle hinlänglich großen x, weil rechterhand beide Summanden $\to 0$ streben für $x\to\infty$, der erste nach (2.11), der zweite nach (2.12). Da schließlich auch $|C|<\varepsilon/4$ ist für alle hinreichend großen x, so ist $|A+B+C|<\varepsilon$ für $x>x_0(\varepsilon)$. Damit ist die Behauptung (2.13) bewiesen.

3. Ein Satz über absolute Inklusion.

Die Gewichtsfunktionen p, q und $p*q=r$ mögen zu $\mathcal{P}$ gehören und es werde wie üblich

$$\int_0^x p(t)\,dt=P(x),\qquad\int_0^x r(t)\,dt=R(x)$$

gesetzt. Dann gilt der

Satz 3. *Jede $|N_p|$-limitierbare Funktion $s(t)\in S$ ist auch $|N_r|$-limitierbar, falls für ein von t unabhängiges M und alle $t\geqq 0$ die Gesamtschwankung*

$$(3.1)\qquad V_{x=t}^{\infty}\frac{1}{R(x)}\int_0^t P(u)\,q(x-u)\,du\leqq M$$

ist. Und wenn überdies für jedes feste $t>0$

$$(3.2)\qquad\frac{1}{R(x)}\int_0^t|q(x-u)|\,du\to 0$$

strebt für $x\to\infty$, so ist auch N_p-lim $s(t)=N_r$-lim $s(t)$, d.h. das Verfahren $|N_p|$ ist in dem Verfahren $|N_r|$ enthalten (in Zeichen: $|N_p|\subseteqq|N_r|$).

Beweis. Wir setzen

$$N_p\big(x,s(t)\big)=\sigma(x),\qquad N_r\big(x,s(t)\big)=\tau(x).$$

Dann ist (vgl. I, § 3) $p*s=P\cdot\sigma$ und also $r*s=q*(P\cdot\sigma)$, d.h.

$$(3.3)\qquad\tau(x)=\int_0^\infty a(x,t)\,\sigma(t)\,dt\qquad\text{mit}\qquad a(x,t)=\frac{q(x-t)\,P(t)}{R(x)}.$$

Für die erste Behauptung des Satzes genügt es nun zu zeigen, daß diese Transformation (3.3) von $\sigma(t)$ in $\tau(x)$ absolut konvergenztreu ist oder daß sie die Bedingungen (2.2) und (2.4) erfüllt. Da aber $\int_0^\infty a(x,t)\,dt=A(x)\equiv 1$

ist, so ist (2.2) gewiß erfüllt. Und im jetzigen Falle ist ersichtlich

$$\varphi(x,t) = 1 - \frac{1}{R(x)} \int_0^t P(u)\, q(x-u)\, du,$$

so daß (2.4) mit (3.1) äquivalent ist.

Für die zweite Behauptung des Satzes ziehen wir die Bemerkungen 2 und 3 zu Satz 1 heran. Es ist jetzt

$$a^*(t) = \lim_{x\to\infty} \int_0^t a(x,u)\, du = \lim_{x\to\infty} \frac{1}{R(x)} \int_0^t P(u)\, q(x-u)\, du,$$

und zwar für jedes $T>0$ gleichmäßig in $0\leq t\leq T$, weil

$$\int_0^t P(u)\, q(x-u)\, du \leq \operatorname*{Max}_{0\leq u\leq T} |P(u)| \cdot \int_0^T |q(x-u)|\, du = o\big(R(x)\big)$$

ist, letzteres nach (3.2). Da schließlich $a=1$ ist, so ist (3.3) absolut regulär.

Der folgende Satz gibt etwas einfachere Bedingungen für dieselbe Inklusion:

Satz 4. *Für die in Satz 3 behauptete Inklusion ist es hinreichend, wenn*

(3.4) $q(t)$ *nicht negativ*

ist und für jedes $u>0$

(3.5) $\frac{q(x-u)}{R(x)} \searrow 0$ *strebt für* $u\leq x \nearrow \infty$.

Beweis. Ersichtlich genügt es zu zeigen, daß (3.1) erfüllt ist. Auf Grund von (3.4) und (3.5) ist aber die in (3.1) auftretende Funktion von x in $x\geq t$ ebenfalls monoton fallend, da die Differenz ihrer Werte für $x=x'$ und $x=x''>x'>t\geq u$

$$= \int_0^t P(u) \left[\frac{q(x'-u)}{R(x')} - \frac{q(x''-u)}{R(x'')}\right] du \geq 0$$

ist. Für festes $t>0$ und $T>t$ ist daher ihre Gesamtschwankung über $t\leq x\leq T$

$$= \frac{1}{R(t)} \int_0^t P(u)\, q(t-u)\, du - \frac{1}{R(T)} \int_0^t P(u)\, q(T-u)\, du.$$

Da hier der Minuend $=1$, der Subtrahend ≥ 0 (aber ≤ 1) ist, ist (3.1) mit $M=1$ erfüllt.

Anwendung. Für

$$p(t) = \frac{t^{k-1}}{\Gamma(k)}, \quad (k>0), \quad \text{und} \quad q(t) = \frac{t^{d-1}}{\Gamma(d)}, \quad (0<d<1),$$

wird, wenn noch $k+d=l$ gesetzt wird,

$$r(t)=p*q=\frac{1}{\Gamma(k)\,\Gamma(d)}\int_0^t (t-u)^{k-1}\,u^{d-1}\,du$$

$$=\frac{t^{l-1}}{\Gamma(k)\,\Gamma(d)}\int_0^1 (1-v)^{k-1}\,v^{d-1}\,dv=\frac{t^{l-1}}{\Gamma(l)}.$$

Die zugehörigen Verfahren N_p und N_r sind ersichtlich die CESÀROschen Verfahren C_k bzw. C_l. Da (3.4) und (3.5) ersichtlich erfüllt sind, letzteres weil abgesehen von konstanten positiven Faktoren

$$\frac{q(x-u)}{R(x)}=\frac{1}{(x-u)^{1-d}\,x^{k+d}}$$

ist und also monoton gegen 0 fällt für wachsendes $x\to\infty$, so ergibt sich aus Satz 4 speziell, daß $|C_k|\leqq|C_l|$ ist für $0\leqq k\leqq l$.

4. Multiplikationssätze.

Wir benutzen die folgenden Bezeichnungen: Wenn die beiden Integrale

$$(4.1)\qquad \int_0^\infty a(t)\,dt, \qquad\qquad (4.2)\qquad \int_0^\infty b(t)\,dt$$

mit $a, b\in S$ vorgelegt sind, setzen wir $a*b=c$ und nennen das Integral

$$(4.3)\qquad \int_0^\infty c(t)\,dt$$

das CAUCHYsche Produkt der beiden ersten. Die Teilintegrale von 0 bis t mögen mit $A(t), B(t), C(t)$ bezeichnet werden und deren NÖRLUND-Mittel, die natürlich jedesmal genauer anzugeben sind, mit $\alpha(x), \beta(x), \gamma(x)$. Die Grenzwerte von diesen sollen schließlich A, B, C genannt werden, wenn sie existieren.

Wenn nun N_p und N_q zwei NÖRLUNDsche Verfahren sind und wenn die Faltung $p*q$ wieder zu $\mathcal{P}$ gehört, so bezeichnen wir sie mit $r(t)$,

$$(4.4)\qquad \int_0^t p(t-u)\,q(u)\,du=r(t),$$

schreiben auch $N_r=N_p*N_q$ (vgl. I, § 5) und setzen endlich die von 0 bis t genommenen Integrale von p, q, r wie bisher beziehentlich $=P(t), Q(t), R(t)$.

Wir können nun den folgenden Satz beweisen, der den klassischen Satz von CESÀRO über die C_1-Summierbarkeit des CAUCHYschen Produktes zweier konvergenter Reihen verallgemeinert:

Satz 5. *Wenn die Verfahren N_p und N_q positiv und regulär sind, wenn* (4.1) *und* (4.2) *beziehentlich N_p- und N_q-summierbar sind zu den Werten A und B, so ist* (4.3) *N_R-summierbar zum Werte $C=A\cdot B$.*

Beweis. Wenn wir $1 * R = \overline{R}$ setzen, so haben wir zu zeigen, daß

$$(4.5) \qquad \bar{\gamma}(x) = \frac{1}{\overline{R}}(R * C) = \frac{1}{\overline{R}(x)} \int_0^x R(x-t)\, C(t)\, dt \to C = A \cdot B$$

strebt für $x \to \infty$. Nun ist aber auf Grund der eingangs eingeführten Bezeichnungen

$$P \cdot \alpha = p * A, \qquad Q \cdot \beta = q * B$$

und also

$$R * C = 1 * p * q * 1 * a * b = p * A * q * B = (P \cdot \alpha) * (Q \cdot \beta)$$

oder

$$(4.6) \qquad \bar{\gamma}(x) = \int_0^x a(x,t)\, \alpha(t)\, \beta(x-t)\, dt$$

mit $a(x,t) = \dfrac{P(t)\, Q(x-t)}{\overline{R}(x)}$ für $0 < t \leq x$ und $= 0$ für $t > x$.

Wir zeigen nun zunächst, daß die Integraltransformation mit diesem Kern $a(x,t)$ regulär ist. Von den drei Bedingungen, die dafür notwendig und hinreichend sind, sind die Norm- und die Zeilenbedingung offenbar erfüllt, weil $a(x,t) \geqq 0$ ist und $\int_0^\infty a(x,t)\, dt = 1$ für alle $x > 0$. Die dritte (Spalten-) Bedingung fordert, daß für jedes feste $c > 0$

$$(4.7) \qquad \frac{1}{\overline{R}(x)} \int_0^c P(t)\, Q(x-t)\, dt \to 0 \quad \text{strebt für} \quad x \to \infty.$$

Nun ist aber

$$\int_0^c P(t)\, Q(x-t)\, dt \leq P(c) \int_0^c Q(x-t)\, dt$$

und

$$\overline{R}(x) = \int_0^x P(x-t)\, Q(t)\, dt \geq \int_0^{x-c} \geq P(c) \int_0^{x-c} Q(t)\, dt.$$

Daher ist die linke Seite von (4.7)

$$\leq \int_0^c Q(x-c)\, dt \Big/ \int_0^{x-c} Q(t)\, dt = \{\overline{Q}(x) - \overline{Q}(x-c)\} / \overline{Q}(x-c),$$

wenn $\int_0^t Q(u)\, du = \overline{Q}(t)$ gesetzt wird. Nun ist aber das Verfahren N_Q nach I, Satz 2, Bemerkung 3 regulär und folglich [vgl. I, (2.5)] strebt der letzte Quotient gegen 0 für $x \to \infty$. Das beweist (4.7) und damit die Regularität des Kernes $a(x,t)$. Vertauscht man p und q, so ergibt sich ebenso die Regularität des Kernes $a(x, x-t)$. Unser Satz ist daher eine unmittelbare Folgerung aus dem folgenden

Lemma. *Sind die Kerne $a(x,t)$ und $a(x, x-t)$ beide positiv und regulär, so folgt aus $\alpha(t) \to A$ und $\beta(t) \to B$, daß die durch* (4.6) *erklärte Funktion*

$$\bar{\gamma}(x) \to A \cdot B \quad \textit{strebt für} \quad x \to \infty.$$

Der Beweis ergibt sich, indem wir (4.6) als eine Integraltransformation mit dem Kern

$$\bar{a}(x,t) = a(x,t)\,\beta(x-t)$$

von $\alpha(t)$ in $\bar{\gamma}(x)$ betrachten. Diese Transformation ist nämlich multiplikativ mit dem Faktor B. Denn die Norm- und die Spaltenbedingung sind offenbar erfüllt, da $a(x,t)$ positiv und regulär und $\beta(t)$ beschränkt ist. Die Zeilenbedingung aber liefert den Faktor B, denn es strebt für $x \to \infty$

$$\int_0^\infty \bar{a}(x,t)\,dt = \int_0^\infty a(x, x-t)\,\beta(t)\,dt \to B,$$

weil $a(x, x-t)$ regulär ist und $\beta(t) \to B$ strebt. Wenn also $\alpha(t) \to A$ strebt, so strebt $\bar{\gamma}(x) \to AB$.

Unter einer geeigneten zusätzlichen Voraussetzung läßt sich sogar die N_r-Summierbarkeit des Produktintegrals beweisen:

Satz 6. *Wenn unter denselben Voraussetzungen wie in Satz 5 noch*

$$k(x) = \frac{1}{q(x)}\,(q * b)_x = O(1) \tag{4.8}$$

ist, so ist

$$N_r\text{-}\int_0^\infty c(t)\,dt = AB. \tag{4.9}$$

Wir zeigen zunächst, daß Satz 5 in diesem Satz 6 enthalten ist. Denn wenn wir in Satz 6 N_q und N_r durch N_Q bzw. N_R ersetzen, so wird aus (4.8) die Bedingung

$$\frac{1}{Q(x)}\,(Q * b)_x = \frac{1}{Q(x)}\,(q * B)_x = \beta(x) = O(1) \tag{4.8'}$$

und Satz 6 besagt dann: Wenn N_p und N_Q positiv und regulär sind und wenn überdies $\beta(x) = O(1)$ ist, so folgt aus $N_p\text{-}\int_0^\infty a(t)\,dt = A$ und $N_Q\text{-}\int_0^\infty b(t)\,dt = B$, daß $N_R\text{-}\int_0^\infty c(t)\,dt = AB$ ist. Und hierin ist offenbar Satz 5 enthalten.

Beweis von Satz 6. Wir haben zu zeigen, daß für $x \to \infty$ das N_r-Mittel

$$\gamma(x) = \frac{1}{R(x)}\,(r * C) = \frac{1}{R}\,(p * q * 1 * a * b) \to AB \tag{4.10}$$

strebt. Nun ist aber, da die Faltung kommutativ und assoziativ ist, $p * 1 * a = p * A = P \cdot \alpha$, $q * b = k \cdot q$, also

$$\gamma(x) = \frac{1}{R(x)} \int_0^x P(t)\,\alpha(t)\,q(x-t)\,k(x-t)\,dt = \int_0^x a(x,t)\,k(x-t)\,\alpha(t)\,dt \tag{4.11}$$

$$\text{mit} \quad a(x,t) = \frac{P(t)\,q(x-t)}{R(x)}.$$

Die Integraltransformation mit diesem Kern $a(x,t)$ ist aber (vgl. den Beweis von Satz 5) wieder regulär. Denn die Norm- und die Zeilenbedingung

sind offenbar erfüllt und die Spaltenbedingung bedeutet hier, daß für jedes $c>0$

$$\frac{1}{R(x)}\int_0^c P(t)\,q(x-t)\,dt\to 0$$

strebt für $x\to\infty$. Da aber der Ausdruck links

$$\leqq \frac{P(c)\int_0^c q(x-t)\,dt}{P(c)\int_0^{x-c} q(t)\,dt} = \frac{Q(x)-Q(x-c)}{Q(x-c)}$$

ist und dieser letzte Quotient, da N_q regulär ist, gegen 0 strebt, so ist auch die vorangehende Relation richtig, also $a(x,t)$ regulär. Ganz ähnlich erkennt man, daß die Transformation mit dem Kern $a(x,t)\cdot k(x-t)$ multiplikativ ist mit dem Faktor B. Denn zunächst ist

$$a(x,t)\,k(x-t) = \frac{P(t)\,q(x-t)}{R(x)}\cdot\frac{(q*b)_{x-t}}{q(x-t)} = \frac{P(t)\cdot(q*b)_{x-t}}{R(x)}$$

und daher wegen (4.8) mit einem passenden $K>0$

$$|a(x,t)\,k(x-t)|\leqq K\,\frac{P(t)\,q(x-t)}{R(x)}.$$

Hieraus aber und auf Grund der für den Kern $a(x,t)$ durchgeführten Rechnung ergibt sich, daß der neue Kern die Norm- und die Spaltenbedingung für seine Regularität erfüllt. An Stelle der Zeilenbedingung gilt aber

$$\int_0^x a(x,t)\,k(x-t)\,dt\to B \quad \text{für} \quad x\to\infty,$$

denn die linke Seite dieser Relation ist (s. oben)

$$= \frac{1}{R(x)}\int_0^x P(t)\,(q*b)_{x-t}\,dt = \frac{1}{R(x)}\,(P*q*b)_x$$

$$= \frac{1}{R(x)}\,(p*q*B)_x = \frac{1}{R(x)}\int_0^x r(x-t)\,B(t)\,dt.$$

Dies letzte Mittel strebt aber $\to B$, weil $N_q\leqq N_r$ gilt (s. I, Satz 2). Also ist (4.11) eine multiplikative Transformation von $\alpha(t)$ in $\gamma(x)$ mit dem Faktor B. Aus $\alpha(t)\to A$ folgt daher $\gamma(x)\to AB$, was den Beweis von Satz 6 vollendet.

War Satz 5 eine Verallgemeinerung des Multiplikationssatzes von CESÀRO, so ist der folgende eine solche des ebenso klassischen Satzes von MERTENS:

Satz 7. *Sind N_p und N_q wieder positiv und regulär und ist* (4.1) *absolut N_p-summierbar zum Werte A und* (4.2) *N_q-summierbar zum Werte B, so ist* (4.3) *N_r-summierbar zum Werte $C=AB$.*

Beweis. Wie in (4.10) haben wir zunächst

$$R \cdot \gamma = r * C = 1 * p * q * a * b,$$

$$1 * (R \cdot \gamma) = p * A * q * B = (P \cdot \alpha) * (Q \cdot \beta)$$

und folglich, wenn wir zur Abkürzung $P \cdot \alpha = f$ und $Q \cdot \beta = g$ setzen, wegen der Stetigkeit von R und γ

$$\gamma(x) = \frac{1}{R(x)} [f * g]'_x. \tag{4.12}$$

Nun sind die vier Funktionen P, α, Q, β und also auch f und g stetig für $x > 0$ und verschwinden für $x \to 0$. Das ist selbstverständlich für P und Q, sowie für A und B. Für $\alpha(x)$ und $x > 0$ folgt die Stetigkeit aus der von $P(x)$ und $\int_0^x p(x-t) A(t)\, dt$ (für die letztere vgl. I, § 3.1). Das letzte Integral ist aber $= \mu \cdot \int_0^x p(t)\, dt = \mu \cdot P(x)$, wenn mit μ ein passender Mittelwert von $A(t)$ in $0 \leqq t \leqq x$ bezeichnet wird. Daher strebt $\alpha(x) = \mu \to A(0) = 0$ für $x \to 0$. Ebenso erkennt man, daß $\beta(x)$ für $x > 0$ stetig ist und für $x \to 0$ verschwindet.

Nach I, § 3.3 ist daher $(f * g)' = f' * g$, wofern f' zur Klasse $\mathcal{R}$ gehört[8]). Es ist aber, da $P(x)$ ein Integral und $\alpha(x)$ nach Voraussetzung von beschränkter Schwankung ist, fast überall

$$f' = (P \cdot \alpha)' = p\alpha + P\alpha'.$$

Also gilt $f' \in \mathcal{R}$, wofern $\alpha' \in \mathcal{R}$ gilt. Das letztere ist aber gewiß der Fall, weil nach Voraussetzung $\int_0^\infty |d\alpha(t)|$ und folglich auch $\int_0^\infty |\alpha'(t)|\, dt$ endlich ist.

Folglich ist $\gamma(x) = \gamma_1(x) + \gamma_2(x)$ mit

$$\begin{cases} \gamma_1(x) = \dfrac{1}{R(x)} \displaystyle\int_0^x p(x-t)\, \alpha(x-t)\, Q(t)\, \beta(t)\, dt \\ \gamma_2(x) = \dfrac{1}{R(x)} \displaystyle\int_0^x P(x-t)\, \alpha'(x-t)\, Q(t)\, \beta(t)\, dt. \end{cases} \tag{4.13}$$

Diese Darstellungen, als Transformationen von $\beta(t)$ in $\gamma_1(x)$ bzw. $\gamma_2(x)$ betrachtet, sind nun nulltreu. Um dies zu erkennen, genügt es zu zeigen, daß die Kerne die Norm- und die Spaltenbedingung in I, § 2 erfüllen. Für den ersten Kern bedeutet die Normbedingung, da $\alpha(t)$ beschränkt ist, daß

$$\int_0^x p(x-t)\, Q(t)\, dt = O(R(x))$$

[8]) In I, § 3 war $\mathcal{R}$ als die Klasse aller derjenigen für $t > 0$ erklärten reellwertigen Funktionen $f(t)$ definiert, die dieselben Voraussetzungen wie die Gewichtsfunktionen aus $\mathcal{P}$ erfüllen, nur daß $\int_0^x f(t)\, dt$ für $x > 0$ nicht immer $\neq 0$ zu sein braucht.

sein soll. Das ist aber trivialerweise der Fall, da die linke Seite $= R(x)$ ist. Für den zweiten Kern bedeutet sie, daß

$$\int_0^x P(x-t)\, Q(t) \cdot |\alpha'(x-t)|\, dt = O(R(x))$$

sein soll. Doch da für $0 \leqq t \leqq x$

$$P(x-t)\, Q(t) = P(x-t) \int_0^t q(u)\, du \leqq \int_0^t P(x-u)\, q(u)\, du \leqq R(x) \tag{4.14}$$

ist, so ist die linke Seite der vorigen Relation in der Tat

$$\leqq R(x) \cdot \int_0^x |\alpha'(t)|\, dt \leqq R(x) \cdot \int_0^\infty |\alpha'(t)|\, dt = O(R(x)).$$

Die Normbedingung ist also für jeden der Kerne erfüllt. Die Spaltenbedingung verlangt für den ersten Kern, daß für jedes $c > 0$

$$\frac{1}{R(x)} \int_0^c Q(t)\, p(x-t)\, dt \to 0$$

strebt. Der Ausdruck zur Linken ist aber, ähnlich wie bei den Schlüssen zu (4.11)

$$\leqq \frac{\int_0^c Q(t)\, p(x-t)\, dt}{\int_{x-c}^{0} p(t)\, Q(x-t)\, dt} \leqq \frac{Q(c)}{Q(c)} \cdot \frac{P(x) - P(x-c)}{P(x-c)}$$

und strebt daher gegen 0. Und für den zweiten Kern hat man bei Benutzung von (4.14)

$$\frac{1}{R(x)} \int_0^c P(x-t)\, Q(t) \cdot |\alpha'(x-t)|\, dt \leqq \int_0^c |\alpha'(x-t)|\, dt = \int_{x-c}^{x} |\alpha'(t)|\, dt \to 0.$$

Die beiden Transformationen (4.13) sind daher nulltreu. Wenn also $\beta(t) \to 0$ strebt, so streben $\gamma_1(x)$ und $\gamma_2(x)$ und folglich auch $\gamma(x) \to 0$. Damit ist Satz 7 für den Spezialfall $B = 0$ bewiesen.

Ist $B \neq 0$, so betrachten wir statt $b(t)$ die Funktion

$$b_1(t) = \begin{cases} b(t) - B & \text{in} \quad 0 \leqq t \leqq 1 \\ b(t) & \text{in} \quad t > 1. \end{cases}$$

Dann ist $\int_0^\infty b_1(t)\, dt$ N_q-summierbar zum Werte 0. Setzen wir also $a * b_1 = c_1$, so ist nach dem eben bewiesenen Spezialfall

$$N_r\text{-}\int_0^\infty c_1(t)\, dt = 0. \tag{4.15}$$

Und hieraus läßt sich die ursprüngliche Behauptung, daß $\int_0^\infty c(t)\, dt$ zum Werte AB N_r-summierbar ist, unschwer ableiten. Es ist nämlich

$$a(x-t)\, b_1(t) = \begin{cases} a(x-t)\,(b(t) - B), & \text{falls} \quad 0 \leqq t \leqq 1 \\ a(x-t)\, b(t), & \text{falls} \quad t > 1 \end{cases}$$

ist, und folglich

$$c_1(x) = \int_0^x a(x-t)\, b_1(t)\, dt =$$

$$= \begin{cases} c(x) - B\int_0^x a(x-t)\, dt = c(x) - B\cdot A(x) & \text{für} \quad 0 \leqq x \leqq 1, \\ c(x) - \int_0^1 a(x-t)\, dt = c(x) - B\big(A(x) - A(x-1)\big) & \text{für} \quad x > 1 \end{cases}$$

oder also für alle $x \geqq 0$

$$c_1(x) = c(x) - B\big(A(x) - A(x-1)\big), \tag{4.16}$$

wenn wir daran festhalten, daß $A(t) = 0$ sein soll für $t < 0$. Setzen wir nun noch $A(t) - A(t-1) = a_1(t)$, so gilt zunächst

$$N_p\text{-}\int_0^\infty a_1(t)\, dt = A. \tag{4.17}$$

Denn es ist

$$\int_0^t a_1(u)\, du = A_1(t) = \int_0^t \big(A(u) - A(u-1)\big)\, du$$

$$= \int_0^t A(u)\, du - \int_0^{t-1} A(u)\, du = \int_{t-1}^t A(u)\, du$$

und folglich ergibt sich (4.17) unmittelbar aus dem folgenden

Hilfssatz. *Wenn das Verfahren N_p positiv und regulär ist und wenn $\int_{t-1}^t s(u)\, du = s_1(t)$ gesetzt wird, so folgt aus N_p-lim $s(t) = s$, daß auch N_p-lim $s_1(t) = s$ ist.*

Beweis. Für die Gewichtsfunktion

$$q_1(t) = \begin{cases} 1 & \text{in} \quad 0 \leqq t \leqq 1 \\ 0 & \text{in} \quad t > 1 \end{cases}$$

ist

$$q_1 * s = \int_{x-1}^x s(t)\, dt = s_1(x), \tag{4.18}$$

$$r_1 = q_1 * p = \int_{x-1}^x p(t)\, dt = P(x) - P(x-1)$$

und

$$R_1(x) = \int_0^x r_1(t)\, dt = \int_{x-1}^x P(t)\, dt. \tag{4.19}$$

Da nun N_{q_1} offenbar positiv und regulär ist und N_p nach Voraussetzung ebenfalls, so besteht nach I, § 2, Satz 2 die Inklusion $N_p \subseteq N_{r_1}$, d.h. aus N_p-lim $s(t) = s$ folgt N_{r_1}-lim $s(t) = s$. Es ist aber $r_1 * s = p * q_1 * s = p * s_1$, letzteres nach (4.18). Es strebt also

$$\frac{1}{R_1(x)} \int_0^x p(x-t)\, s_1(t)\, dt \to s.$$

Nach (4.19) ist aber $R_1(x) = P(\xi)$ für ein geeignetes ξ zwischen $x-1$ und x und also asymptotisch gleich $P(x)$. Die letzte Beziehung liefert daher die Behauptung des Hilfssatzes und vollendet dadurch den Beweis von (4.17).

Nach (4.16) ist nun

$$c(t) = c_1(t) + B\, a_1(t)$$

und folglich unter Berücksichtigung von (4.15)

$$N_r\text{-}\int_0^\infty c(t)\,dt = B \cdot \left\{ N_r\text{-}\int_0^\infty a_1(t)\,dt \right\}.$$

Wegen (4.17) und $N_p \subseteq N_r$ ist endlich der Wert der in der letzten Beziehung rechts stehenden Klammer $= A$, so daß diese Beziehung die Behauptung des Satzes 7 darstellt.

Literatur.

[1] KNOPP, K., and B. VANDERBURG: Functional NÖRLUND-Methods. I. Rend. Circ. Mat. Palermo (2) **4**, Fasc. 1 (1955). — [2] MEARS, FL. M.: Some Multiplication Theorems for NÖRLUND Means. Bull. Amer. Math. Soc. (2) **41**, 875—880 (1935). — [3] MEARS, FL. M.: Absolute Regularity and the NÖRLUND Mean. Ann. of Math. (2) **38**, 594—601 (1937). — [4] MEARS, FL. M.: The Inverse NÖRLUND Mean. Ann. of Math. (2) **44**, 401—410 (1943). — [5] MEARS, FL. M.: NÖRLUND Summability of CAUCHY Products. Ann. of Math. (2) **46**, 563—566 (1945). — [6] SUNOUCHI, G., and T. TSUCHIKURA: Absolute Regularity for Convergent Integrals. Tôhoku Math. J. (2) **4**, 153—156 (1952).

University of Cincinnati.

Universität Tübingen.

(Eingegangen am 22. Dezember 1954.)

Dieudonné, J.
Math. Zeitschr. Bd. 63, S. 53—75 (1955)

Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$. III.

A la mémoire du grand algébriste I. Schur.

Par

Jean Dieudonné.

1. Introduction.

Ce travail fait suite à plusieurs articles antérieurs ([2], [3], [5]), dont nous conservons en principe la terminologie et les notations; il se rattache plus particulièrement à [5], étant exclusivement consacré à l'étude de certains groupes de Lie *abéliens*. Nous y développons notamment une méthode qui permet de démontrer l'*existence* de types très nombreux de groupes de Lie abéliens, deux à deux non isomorphes; parmi ces types figurent les groupes à un paramètre I_r dont l'hyperalgèbre est définie dans [5], et dont nous n'avions pu alors établir l'existence. La méthode dont il s'agit repose sur une étude préalable des *groupes de type hyperexponentiel*, dont la structure est particulièrement simple et qui jouent le rôle d'une sorte de «revêtement universel» pour les groupes que nous étudions; en outre, leur théorie présente un intérêt par elle-même, car elle permet, par exemple, de faire l'étude complète des *groupes additifs de Witt*, du point de vue «infinitésimal».

2. Séries formelles à une infinité de variables graduées.

Il s'avère nécessaire, pour notre objet, de généraliser la notion de groupe de Lie formel, en l'étendant à des séries formelles à une infinité de variables. De façon précise, considérons une suite infinie d'indéterminées $x_0, x_1, \ldots, x_i, \ldots$, et supposons qu'à chacune d'elles soit attaché un entier $\pi_i > 0$, son *poids*, soumis à la condition que *la suite* (π_i) *est strictement croissante et tend vers* $+\infty$; on forme comme d'ordinaire des *monômes*, produits de puissances d'un nombre *fini* des x_i, le poids du produit de deux monômes étant la somme de leurs poids. En raison de l'hypothèse sur la suite (π_i), il n'y a qu'un nombre *fini* de monômes de poids inférieur à un entier donné. Pour un système $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots)$ d'entiers $\geqq 0$, nuls sauf un nombre fini d'entre eux, nous désignerons encore par x_α le monôme $\prod_{i=0}^{\infty} x_i^{\alpha_i}$; nous désignerons par $\pi(\alpha)$ le poids $\sum_i \alpha_i \pi_i$ de ce monôme; sa hauteur $h(\alpha)$ sera encore le plus petit nombre h tel que $\alpha_i < p^{h+1}$ pour tout indice i; les notations $|\alpha|$ et $\alpha!$ auront le même sens que dans [2, n° 5].

Cela étant, les séries formelles en $x_0, \ldots, x_i, \ldots$ à coefficients dans un corps K sont encore les combinaisons linéaires infinies $\sum_\alpha c_\alpha x_\alpha$, où $c_\alpha \in K$, que l'on ajoute et multiplie comme d'ordinaire [les poids n'interviennent pas ici, mais seulement le fait que pour tout α, il n'y a qu'un nombre fini de couples (β, γ) tels que $\alpha = \beta + \gamma$]. La substitution aux x_i de séries formelles sans terme constant, par rapport à des variables graduées $y_0, y_1, \ldots, y_i, \ldots$, ne peut plus ici se faire sans restriction; elle a toutefois toujours un sens si le *poids minimum* de la série f_i que l'on substitue à x_i (c'est-à-dire le plus petit des poids de ses termes $\neq 0$) *tend vers* $+\infty$ *avec* i: il n'y a alors qu'un nombre fini de monômes x_α qui, lorsqu'on y substitue les f_i aux x_i, fournissent des termes de poids donné.

Nous pouvons alors étendre la définition d'un groupe de Lie formel; il doit y être sous-entendu que, dans tous les systèmes d'indéterminées (x_i), (y_i), etc. que l'on considère, le poids de la variable d'indice i est toujours le *même* entier π_i. La donnée d'un groupe de Lie formel G à une infinité de variables (ou «de dimension infinie») consiste alors en la donnée d'une suite (φ_i) de séries formelles par rapport à deux systèmes (x_i), (y_i) de variables graduées, avec les conditions suivantes: 1° chaque φ_i est de poids minimum π_i; 2° $\varphi_i(x_0, \ldots, x_i, \ldots; 0, \ldots, 0, \ldots) = x_i$, $\varphi_i(0, \ldots, 0, \ldots; y_0, \ldots, y_i, \ldots) = y_i$; 3° $\varphi_i\big(\varphi(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{z}\big) = \varphi_i\big(\boldsymbol{x}, \varphi(\boldsymbol{y}, \boldsymbol{z})\big)$ pour tout i, avec les conventions habituelles (les substitutions ayant un sens grâce à la condition 1°). On a alors

$$\varphi_i(\boldsymbol{x}, \boldsymbol{y}) = x_i + y_i + \psi_i(\boldsymbol{x}, \boldsymbol{y})$$

où ψ_i est une série formelle dont chaque monôme contient au moins un x_j et au moins un y_j, et qui est de poids minimum $\geq \pi_i$. On peut encore démontrer l'existence d'un «inverse», c'est-à-dire de séries $\vartheta_i(\boldsymbol{x}) = -x_i + \omega_i(\boldsymbol{x})$, où ω_i est de poids minimum $> \pi_i$, telles que $\varphi_i\big(\boldsymbol{x}, \vartheta(\boldsymbol{x})\big) = 0$ pour tout i: il faut ici remplacer la récurrence classique sur le degré des termes des ω_i par une récurrence sur les poids.

Un *homomorphisme* d'un groupe formel G dans un groupe formel G', lorsque G et G' sont de dimension infinie, est un système $\boldsymbol{u} = (u_i)$ de séries formelles $u_i(x_0, \ldots, x_j, \ldots)$ telles que le poids minimum de u_i *tende vers* $+\infty$ *avec* i, et que l'on ait $\varphi'\big(\boldsymbol{u}(\boldsymbol{x}), \boldsymbol{u}(\boldsymbol{y})\big) = \boldsymbol{u}\big(\varphi(\boldsymbol{x}, \boldsymbol{y})\big)$. Les notions d'isomorphisme et de «changement de variables» se définissent de même. On a aussi des définitions analogues lorsque l'un des deux groupes G, G' est de dimension finie (lorsque c'est G', il n'y a naturellement plus aucune restriction sur les u_i).

3. Groupes récurrents.

Soit $\mathfrak{o}$ l'anneau des séries formelles par rapport à une infinité de variables graduées $x_0, \ldots, x_i, \ldots$. La notion générale de *semi-dérivation* se définit comme dans [2], ainsi que les semi-dérivations D_α: si on observe que $D_\alpha x_\beta$ est nul ou un monôme de poids $\pi(\beta) - \pi(\alpha)$, on peut en effet définir $D_\alpha\big(\sum_\beta c_\beta x_\beta\big)$ comme $\sum_\beta c_\beta (D_\alpha x_\beta)$ (avec $c_\beta \in K$), cette série formelle ayant un sens. Un

opérateur différentiel D sur $\mathfrak{o}$ pourra cette fois être défini comme une somme infinie $\sum_\alpha v_\alpha D_\alpha$, où les v_α appartiennent à un anneau de séries formelles contenant $\mathfrak{o}$, pourvu que le poids minimum de v_α *augmente indéfiniment avec* $\pi(\alpha)$: la série $\sum_\alpha v_\alpha D_\alpha f$ est en effet définie dans ce cas pour toute $f \in \mathfrak{o}$; en particulier, on a la *formule de Taylor*

$$f(\boldsymbol{x}+\boldsymbol{z}) = \sum_\alpha \frac{1}{\alpha!} z_\alpha D_\alpha f$$

puisque cette formule est vraie pour tout monôme et que z_α est de poids $\pi(\alpha)$.

Nous ne rechercherons pas les conditions les plus générales dans lesquelles une «théorie de Lie» est possible pour les groupes de dimension infinie, et allons nous borner aux seuls groupes qui interviendront par la suite, que nous appellerons *groupes récurrents*. Un tel groupe G est caractérisé par la condition que, pour tout indice i, $\varphi_i(x_0, \ldots, x_j, \ldots; x_0, \ldots, y_j, \ldots)$ *ne contient pas les variables* x_j, y_j *d'indice* $j > i$; cela entraîne que les équations $x_0 = \cdots = x_{i-1} = 0$ définissent un *sous-groupe* (récurrent) G_i' de G; d'autre part, les i séries formelles $\varphi_j(x_0, \ldots, x_{i-1}, 0, \ldots; y_0, \ldots, y_{i-1}, 0, \ldots)$ $(0 \leq j \leq i-1)$ définissent un groupe G_i de dimension i, et le système $\boldsymbol{u}^{(i)} = (u_j^{(i)})$ de séries formelles $u_j^{(i)}(\boldsymbol{x}) = x_j$ pour $j < i$ définit un homomorphisme de G sur G_i, de noyau G_i'. Etant donnés deux groupes récurrents $G, \overline{G}$, un *homomorphisme récurrent* $\boldsymbol{u}$ de G dans $\overline{G}$ est par définition tel que $u_i(\boldsymbol{x})$ *ne contienne que les variables* x_j *d'indice* $j \leq i$; sa restriction à G_i' est un homomorphisme (récurrent) de ce sous-groupe sur $\overline{G}_i'$ pour tout i, et les i premières séries u_j définissent un homomorphisme de G_i dans $\overline{G}_i$.

La définition des opérateurs de translation $L_{\boldsymbol{y}}$ et $R_{\boldsymbol{y}}$, et des opérateurs différentiels *invariants à gauche* se fait comme dans [2, n^{os} 6 et 7] pour un groupe quelconque de dimension infinie. Si Z_α est l'opérateur différentiel invariant tel que $Z_\alpha(\boldsymbol{e}) = \frac{1}{\alpha!} D_\alpha(\boldsymbol{e})$, $Z_\alpha f$ est le coefficient de y_α dans le développement de $f(\boldsymbol{x}\,\boldsymbol{y})$ suivant les puissances des y_i, et on a

$$Z_\alpha = \frac{(1+w_\alpha)}{\alpha!} D_\alpha + \sum_\beta w_\beta D_\beta \tag{1}$$

où w_α et les w_β sont des séries sans terme constant, et la sommation est étendue aux β tels que $|\beta| \leq |\alpha|$ et $\beta \neq \alpha$; cette sommation est ici infinie, mais garde un sens car, en vertu de la formule de Taylor et de l'hypothèse sur les φ_i, le poids minimum de w_β augmente indéfiniment avec $\pi(\beta)$.

Supposons désormais que le groupe G soit *récurrent*; alors on a en particulier

$$X_{hi} = (1 + w_{hi}^{(i)}) D_{hi} + \sum_{j>i} w_{hj}^{(i)} D_{hj} + \sum_\beta w_\beta^{(i)} D_\beta \tag{2}$$

où toutes les séries $w^{(i)}$ sont sans terme constant, les $w_{hj}^{(i)}$ $(j \geq i)$ appartiennent à $\mathfrak{o}_h$, et les D_β sont des semi-dérivations *spéciales* de hauteur h. En outre, le poids minimum de $w_{hj}^{(i)}$ est $\geq p^h(\pi_j - \pi_i) \geq p^h(j-i)$. Avec ces semi-dérivations invariantes, on forme les «monômes» X_α, qu'il est plus

commode ici de prendre dans l'ordre inverse de celui choisi dans [2, n° 8]: si $\alpha = (\alpha_0, \ldots, \alpha_i, 0, 0, \ldots)$ et $\alpha_j = \sum_{h=0}^{r} v_{hj} p^h$, on pose

$$X_\alpha = X_{r0}^{v_{r0}} \ldots X_{ri}^{v_{ri}} X_{r-1,0}^{v_{r-1,0}} \ldots X_{r-1,i}^{v_{r-1,i}} \ldots X_{00}^{v_{00}} \ldots X_{0i}^{v_{0i}}.$$

Le point essentiel consiste à préciser les propriétés d'«indépendance linéaire» des D_α et des X_α. En premier lieu, une relation

$$\sum_\mu u_\mu D_\mu = 0 \tag{3}$$

où les u_μ appartiennent à un anneau de séries formelles contenant $\mathfrak{o}$, et sont telles que le poids minimum de u_μ augmente indéfiniment avec $\pi(\mu)$, entraîne $u_\mu = 0$ pour tout μ. Pour le voir, on raisonne par l'absurde et on considère dans (3) les termes (en nombre fini) de poids minimum ϱ dans *tous* les u_μ; si ces termes proviennent des séries $u_\alpha, u_\beta, \ldots, u_\lambda$ et si on désigne leur somme par $w_\alpha, w_\beta, \ldots, w_\lambda$ respectivement (dans chacune de ces séries), en écrivant que l'opérateur du premier membre de (3), appliqué successivement aux monômes $x_\alpha, x_\beta, \ldots, x_\lambda$, donne 0, on obtient des séries dont les termes de poids ϱ se réduisent respectivement à $w_\alpha, w_\beta, \ldots, w_\lambda$; ces séries étant nulles, il était contradictoire de supposer $w_\alpha, \ldots, w_\lambda$ non nuls, c'est-à-dire les u_μ non tous nuls.

Remarquons ensuite qu'on peut tirer des équations (2) les formules

$$D_{hi} = (1 + u_{hi}^{(i)}) \Big(X_{hi} - \sum_\beta w_\beta^{(i)} D_\beta \Big) + \sum_{j>i} u_{hj}^{(i)} \Big(X_{hj} - \sum_\beta w_\beta^{(j)} D_\beta \Big) \tag{4}$$

où les $u_{hj}^{(i)}$ $(j \geqq i)$ sont des séries sans terme constant, appartenant à $\mathfrak{o}_h$, et telles que le poids minimum de $u_{hj}^{(i)}$ augmente indéfiniment avec j. En effet, la détermination des $u_{hj}^{(i)}$ se fait par le système récurrent

$$(1 + u_{hi}^{(i)})(1 + w_{hi}^{(i)}) = 1$$

$$(1 + u_{hi}^{(i)}) w_{hj}^{(i)} + \sum_{k=i+1}^{j-1} u_{hk}^{(i)} w_{hj}^{(k)} + (1 + w_{hj}^{(j)}) u_{hj}^{(i)} = 0 \quad \text{pour} \quad j > i$$

d'où l'en déduit, par récurrence sur $j - i$, que le poids minimum de $u_{hj}^{(i)}$ est $\geqq p^h(j-i)$, ce qui établit nos assertions.

L'étape suivante consiste à montrer que pour tout α, on peut écrire

$$D_\alpha = v_\alpha X_\alpha + \sum_\beta v_\beta X_\beta \tag{5}$$

où la sommation est étendue aux β tels que $|\beta| \leqq |\alpha|$, le poids minimum de v_β augmentant indéfiniment avec $\pi(\beta)$; en outre, le terme constant de v_α est égal à 1 (un nombre *fini* d'autres v_β peuvent avoir un terme constant $\neq 0$). Ceci se fait par récurrence sur la hauteur $h(\alpha)$, comme dans [2, n^{os} 8 et 10] (la nouvelle convention adoptée pour les X_α modifiant un peu la démonstration). On a d'ailleurs une formule analogue à (5), avec les mêmes propriétés pour

les coefficients, en y intervertissant les rôles des X_α et des D_α:

$$X_\alpha = t_\alpha D_\alpha + \sum_\beta t_\beta D_\beta. \tag{6}$$

Cette formule est d'ailleurs beaucoup plus élémentaire que (5), et découle aussitôt de (2) et de la définition de X_α, en faisant usage des «formules de Leibniz» pour les D_{hi}. On peut la préciser de la façon suivante. Définissons sur l'ensemble des indices α un ordre total qu'on pourrait appeler «ordre lexicographique inverse par étages»: si $\alpha = (\alpha_0, \ldots, \alpha_i, \ldots)$, $\beta = (\beta_0, \ldots, \beta_i, \ldots)$, α est «avant» β si: 1° ou bien $h(\alpha) > h(\beta)$; 2°, ou bien $h(\alpha) = h(\beta)$, et, si on pose $\alpha_i = \sum_h \mu_{hi} p^h$, $\beta_i = \sum_h \nu_{hi} p^h$, pour le plus grand $h \leqq h(\alpha)$ tel que $\mu_{hi} \neq \nu_{hi}$ pour un i au moins, on a $\mu_{hi} > \nu_{hi}$ pour le plus petit indice i pour lequel $\mu_{hi} \neq \nu_{hi}$. On voit alors par récurrence sur $|\alpha|$, en tenant compte de (2), que dans la formule (6), t_α est précisément le «premier» terme non nul, pour l'ordre qui vient d'être défini.

Cette dernière remarque permet enfin de prouver l'indépendance linéaire des X_α, au même sens que pour les D_α. Supposons qu'une relation

$$\sum_\mu u_\mu X_\mu = 0 \tag{7}$$

ait ses coefficients u_μ non nuls; soient encore $w_\alpha, \ldots, w_\lambda$ les termes de poids minimum ϱ dans les u_μ, en nombre fini, et supposons $\alpha, \beta, \ldots, \lambda$ ordonnés suivant l'ordre qui vient d'être décrit. Si on remplace les X_μ dans (7) par leurs expressions (6), on a vu que dans la relation de type (3) entre les D_μ que l'on obtient, tous les coefficients doivent être nuls. Or, les termes de poids minimum dans le coefficient de D_α ont exactement pour somme w_α, puisque α est le «premier» des indices $\alpha, \beta, \ldots, \lambda$ (on se souviendra que le terme constant de t_α est 1). On aurait donc $w_\alpha = 0$, ce qui contredit l'hypothèse. Un raisonnement analogue, mais plus simple, basé sur la formule (1), démontre l'indépendance linéaire des Z_α.

A partir de ces résultats, on peut développer le reste de la théorie exactement comme dans le cas fini. Nous en utiliserons donc librement les résultats, et nous nous bornerons à ajouter quelque remarques tenant à la définition particulière des groupes récurrents. Si dans la formule (1) on remplace les D_μ en utilisant (6) on obtient une relation

$$Z_\alpha = \sum_\mu u_\mu X_\mu \tag{8}$$

et comme Z_α et les X_μ sont des semi-dérivations invariantes, le raisonnement du th. 2 de [2] montre, compte tenu de l'indépendance linéaire des X_μ, que les u_μ sont des *constantes*. Il n'y a par suite qu'un nombre *fini* de ces constantes qui sont $\neq 0$: d'après la formule (1), elles correspondent exactement aux indices μ tels que, dans (5), la série v_μ ait un terme constant $\neq 0$; en particulier, le coefficient de X_α dans (8) est $\frac{1}{\alpha!}$, et les autres coefficients u_μ non nuls correspondent à des indices tels que $|\mu| \leqq |\alpha|$. Le fait que G est

un groupe récurrent entraîne en outre que les X_α dont les indices α ont toutes leurs composantes d'indice $k<i$ nulles, forment la base d'une *sous-algèbre* $\mathfrak{G}'_i$ de l'hyperalgèbre $\mathfrak{G}$ de G, savoir l'hyperalgèbre du sous-groupe G'_i [2, n° 18]. Si, dans la formule (8), α a toutes ses composantes d'indice $<i$ nulles, les seuls indices μ tels que $u_\mu \neq 0$ sont donc des indices ayant la même propriété.

Notons enfin que si $\boldsymbol{u}^{(i)}$ est l'homomorphisme «canonique» de G sur G_i défini ci-dessus, l'homomorphisme dérivé $(\boldsymbol{u}^{(i)})'$ a pour noyau l'*idéal* $\mathfrak{m}_i$ de l'hyperalgèbre de G ayant pour base les X_α (ou les Z_α) tels qu'*une au moins des composantes de* α *d'indice* $\geqq i$ *soit non nulle*; cela résulte de la formule (13) de [5] et des formules

$$(\boldsymbol{u}^{(i)})'(X_{hj}) = X_{hj} \quad \text{pour} \quad j \leqq i$$
$$(\boldsymbol{u}^{(i)})'(X_{hj}) = 0 \quad \text{pour} \quad j > i.$$

Cette dernière remarque entraîne la propriété suivante: si $\alpha = (\alpha_0, \ldots, \alpha_i, 0, 0, \ldots)$, la série $P_\alpha(\boldsymbol{y})$, coefficient de $X_\alpha f$ dans $R_{\boldsymbol{y}} f$, *ne contient que les variables* x_j *d'indice* $j \leqq i$; il suffit d'appliquer a $\boldsymbol{u}^{(i)}$ les formules (38) et (39) de [2]. Le procédé standard pour obtenir une loi de composition canonique [3] conduit donc à un groupe récurrent (à loi canonique) $\overline{G}$, et définit un isomorphisme récurrent de G sur $\overline{G}$.

4. Les théorèmes d'existence.

Dans [5], nous avons démontré deux propositions générales (prop. 1 et 2) permettant de «remonter» d'une hyperalgèbre de Lie (resp. d'un homomorphisme d'hyperalgèbres) à un groupe (resp. un homomorphisme de groupes). Nous allons leur donner une forme légèrement différente, mieux adaptée aux applications que nous avons en vue. Nous commencerons par nous borner aux groupes de dimension finie.

La base (Z_α) de l'hyperalgèbre de Lie $\mathfrak{G}$ d'un groupe G de dimension n vérifiant la table de multiplication

$$Z_\alpha Z_\beta = \sum_\gamma d_{\alpha\beta\gamma} Z_\gamma, \tag{9}$$

on a les formules

$$Z_\beta x_\alpha = \sum_\gamma d_{\gamma\beta\alpha} x_\gamma \tag{10}$$

et la loi de groupe G s'exprime par les formules

$$\varphi_i(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\alpha,\beta} d_{\alpha\beta\varepsilon_i} x_\alpha y_\beta \qquad (1 \leq i \leq n). \tag{11}$$

Dans la prop. 1 de [5], on a montré que si, réciproquement, on se donne (abstraitement) une algèbre $\mathfrak{G}$ ayant la table de multiplication (9), elle est l'hyperalgèbre de Lie d'un groupe G *défini* par (11), et on a les relations (10), pourvu que l'on ait les trois conditions suivantes: 1° $d_{0\alpha\beta} = d_{\alpha 0\beta} = \delta_{\alpha\beta}$, $d_{\alpha 0\beta} = 0$ sauf pour $\alpha = \beta = 0$; 2° pour tout $r > 0$, les Z_α de hauteur $h(\alpha) \leqq r$ forment la

base d'une sous-algèbre de $\mathfrak{G}$; 3° les $d_{\alpha\beta\gamma}$ vérifient les relations

$$d_{\alpha,\beta,\lambda+\mu} = \sum_{0\leq\gamma\leq\alpha,\,0\leq\delta\leq\beta} d_{\gamma\delta\lambda}\, d_{\alpha-\gamma,\beta-\delta,\mu} \tag{12}$$

pour tous les indices $\alpha, \beta, \lambda, \mu$. Or, cette dernière condition peut se transformer de la façon suivante: considérons l'algèbre $\mathfrak{G}\otimes\mathfrak{G}$, *produit tensoriel* de $\mathfrak{G}$ par elle-même, qui a pour base les produits $Z_\alpha\otimes Z_\beta$. Pour tout indice α, posons

$$Z^0_\alpha = \sum_{0\leq\gamma\leq\alpha} Z_\gamma \otimes Z_{\alpha-\gamma}. \tag{13}$$

On a, d'après (9),

$$\begin{aligned} Z^0_\alpha Z^0_\beta &= \sum_{\gamma,\delta} (Z_\gamma\otimes Z_{\alpha-\gamma})\,(Z_\delta\otimes Z_{\beta-\delta}) \\ &= \sum_{\gamma,\delta} \Big(\sum_\lambda d_{\gamma\delta\lambda} Z_\lambda\Big)\otimes\Big(\sum_\mu d_{\alpha-\gamma,\beta-\delta,\mu} Z_\mu\Big) \\ &= \sum_{\gamma,\delta,\lambda,\mu} d_{\gamma\delta\lambda}\, d_{\alpha-\gamma,\beta-\delta,\mu}\, Z_\lambda\otimes Z_\mu \end{aligned}$$

et par suite les relations (12) sont *équivalentes* aux relations

$$Z^0_\alpha Z^0_\beta = \sum_\gamma d_{\alpha\beta\gamma} Z^0_\gamma. \tag{14}$$

Ces dernières entraînent que les Z^0_α engendrent une *sous-algèbre* $\mathfrak{G}^0$ de $\mathfrak{G}\otimes\mathfrak{G}$ (d'ailleurs isomorphe à $\mathfrak{G}$)[1]. Mais inversement, si cette propriété est vérifiée, toutes les sommes du second membre de (12) qui correspondent à la même valeur ϱ de $\lambda+\mu$ sont égales; en particulier, faisant $\mu=0$, on voit que cette valeur n'est autre que $d_{\alpha\beta\varrho}$ (puisque $d_{\alpha\beta 0}=0$ sauf pour $\alpha=\beta=0$), et on a donc les relations (12). En résumé:

Proposition 1. *Pour qu'une algèbre associative $\mathfrak{G}$, ayant une base (Z_α) dont les indices parcourent $\boldsymbol{N}^n$, et telle que Z_0 soit l'élément unité de $\mathfrak{G}$, soit l'hyperalgèbre d'un groupe de Lie de dimension n, il faut et il suffit que*: 1° *les Z_α d'indice $\alpha\neq 0$ forment la base d'une sous-algèbre de $\mathfrak{G}$*; 2° *pour tout $r\geqq 0$, les Z_α de hauteur $h(\alpha)\leq r$ forment la base d'une sous-algèbre de $\mathfrak{G}$*; 3° *les Z^0_α définis par (13) forment la base d'une sous-algèbre $\mathfrak{G}^0$ de $\mathfrak{G}\otimes\mathfrak{G}$.*

Un calcul tout à fait analogue permet de transformer la prop. 2 de [5] en la suivante:

Proposition 2. *Pour qu'un homomorphisme $\boldsymbol{u}'$ de l'hyperalgèbre $\mathfrak{G}$ de G dans l'hyperalgèbre $\overline{\mathfrak{G}}$ de $\overline{G}$, transformant l'élément unité en élément unité, soit dérivé d'un homomorphisme de G dans $\overline{G}$, il faut et il suffit que $\boldsymbol{u}'(\mathfrak{g}_r)\subset\overline{\mathfrak{g}}_r$ pour tout $r\geqq 0$ et que $\boldsymbol{u}'\otimes\boldsymbol{u}'$ transforme $\mathfrak{G}^0$ en une sous-algèbre de $\overline{\mathfrak{G}}^0$.*

Les deux propositions précédentes s'étendent sans peine aux groupes récurrents de dimension infinie. Des formules (1), et en se souvenant que $Z_\alpha(\boldsymbol{e}) = \frac{1}{\alpha!} D_\alpha(\boldsymbol{e})$ par définition, on tire aisément les relations (9), où les

[1] Il est facile de voir que l'isomorphisme de $\mathfrak{G}$ sur $\mathfrak{G}^0$ qui fait correspondre Z^0_α à Z_α n'est autre que le *dérivé* de l'isomorphisme «diagonal» $\boldsymbol{x}\to(\boldsymbol{x},\boldsymbol{x})$ de G dans le groupe produit $G\times G$.

$d_{\alpha\beta\gamma} \neq 0$ sont en nombre *fini* et correspondent à des indices γ tels que $|\gamma| \leq |\alpha| + |\beta|$: de façon plus précise, il n'y a qu'un seul de ces indices pour lequel les deux membres de l'inégalité précédente sont égaux, savoir $\gamma = \alpha + \beta$, avec $d_{\alpha,\beta,\alpha+\beta} = \frac{(\alpha+\beta)!}{\alpha!\,\beta!}$; les formules (10) et (11) se démontrent alors comme dans [2, n° 11]. L'hypothèse que G est récurrent s'exprime, en vertu de (11), par les conditions suivantes: on a $d_{\alpha\beta\varepsilon_i} = 0$ si $\pi(\alpha+\beta) < \pi_i$, ou si une des composantes de α ou de β, d'indice $> i$, n'est pas nulle. Inversement, ces conditions, jointes à celles de la proposition 1, assurent que l'algèbre $\mathfrak{G}$ est l'hyperalgèbre d'un groupe récurrent.

Avant de passer aux applications des propositions précédentes, observons en passant que les raisonnements qui y conduisent s'appliquent aussi bien aux groupes de Lie formels sur un corps de caractéristique 0, et conduisent alors aussitôt au «troisième théorème de Lie» pour ces groupes: une fois que l'on sait (th. de Birkhoff-Witt) que l'algèbre enveloppante $\mathfrak{G}$ d'une algèbre de Lie $\mathfrak{g}$ a pour base les X_α, l'hypothèse que $\mathfrak{g}$ est une algèbre de Lie entraîne que les crochets $[X_i^0, X_j^0]$ s'expriment comme combinaisons linéaires des X_k^0, et la prop. 1 montre alors que $\mathfrak{g}$ est l'algèbre de Lie d'un groupe formel.

5. Le groupe hyperexponentiel.

Dans tout ce qui suit, on prend $\pi_i = p^i$. La *série hyperexponentielle* sur un corps K de caractéristique $p > 0$ est une série formelle $\mathrm{Hex}(x_0, x_1, \ldots, x_i, \ldots)$ par rapport aux variables graduées x_i, qui se définit comme suit [*1*, n° 4]: on considère, sur le corps $\boldsymbol{Q}$ des rationnels, la série

$$\exp(u_0 + u_1 + \cdots + u_i + \cdots) \tag{15}$$

où les u_i sont des indéterminées, u_i étant affectée du poids p^i; les termes de poids p^i dans cette série ont pour somme un polynôme de la forme $u_i + F_i(u_0, \ldots, u_{i-1})$. Il est immédiat, par récurrence, qu'il existe des polynômes $G_i(x_0, \ldots, x_{i-1})$ à coefficients rationnels, tels que G_i soit isobare de poids p^i, et que le système d'équations $x_i = u_i + F_i(u_0, \ldots, u_{i-1})$ $(i = 0, 1, 2, \ldots)$ se résolve par $u_i = x_i + G_i(x_0, \ldots, x_{i-1})$. Substituant aux u_i ces polynômes dans la série (15), on obtient une série à coefficients rationnels par rapport aux x_i, et on démontre [*1*, lemme 2] que les coefficients de cette série sont des *entiers p-adiques*. Remplaçant ces coefficients par leurs restes mod. p dans le corps $\boldsymbol{F}_p$, on obtient une série formelle sur ce corps:

$$\mathrm{Hex}(x_0, x_1, \ldots, x_i, \ldots) = \sum_{k=0}^{\infty} E_k(x_0, \ldots, x_i, \ldots) \tag{16}$$

où on désigne par E_k le polynôme somme des termes de poids k. La propriété fondamentale de cette série est que l'on a

$$\mathrm{Hex}(x_0, x_1, \ldots, x_i, \ldots)\,\mathrm{Hex}(y_0, y_1, \ldots, y_i, \ldots) = \mathrm{Hex}(z_0, z_1, \ldots, z_i, \ldots) \tag{17}$$

quand on fait dans cette relation la substitution

$$z_i = \varphi_i(x_0, \ldots, x_i; y_0, \ldots, y_i) = x_i + y_i + \sum_{1<k<p^i} E_k(\boldsymbol{x})\, E_{p^i-k}(\boldsymbol{y}) \tag{18}$$

et par suite les formules (18) définissent un groupe récurrent abélien, que nous appellerons *groupe hyperexponentiel* et noterons H ou H_∞. Ses propriétés fondamentales sont résumées dans l'énoncé suivant:

Proposition 3. 1° *Le système des séries formelles* $\boldsymbol{t} = (t_i)$ *telles que*

$$t_0(\boldsymbol{x}) = 0, \qquad t_i(\boldsymbol{x}) = x_{i-1} \qquad (i = 1, 2, \ldots)$$

définit un homomorphisme de H *dans lui-même, appliquant* H *sur* H'_0.

2° *L'hyperalgèbre de* H *est définie par les relations*

$$X_{hi}^p = X_{h,i+1} \qquad (h = 0, 1, \ldots;\ i = 0, 1, \ldots), \tag{19}$$

et la loi de groupe de H *est canonique.*

1° Il est immédiat, par définition, que l'on a

$$F_i(0, u_0, \ldots, u_{i-2}) = F_{i-1}(u_0, \ldots, u_{i-2})$$

d'où on tire, par récurrence sur i

$$G_i(0, x_0, \ldots, x_{i-2}) = G_{i-1}(x_0, \ldots, x_{i-2})$$

et par suite

$$\mathrm{Hex}(0, x_0, x_1, \ldots) = \mathrm{Hex}(x_0, x_1, \ldots)$$

ce qui, en vertu de (17) et (18), donne

$$\varphi_i(0, x_0, \ldots, x_{i-1}; 0, y_0, \ldots, y_{i-1}) = \varphi_{i-1}(x_0, \ldots, x_{i-1}; y_0, \ldots, y_{i-1}) \tag{21}$$

et démontre notre assertion. La définition de l'homomorphisme $\boldsymbol{t}$ prouve en outre [2, n° 12] que l'on a

$$\boldsymbol{t}'(X_{hi}) = X_{h,i+1} \qquad (h = 0, 1, \ldots;\ i = 0, 1, \ldots). \tag{22}$$

2° En raison des formules (22), pour démontrer les relations (19), il suffit de considérer le cas $i = 0$. Commençons par calculer X_{00}^p; par définition, on a

$$X_{00} = D_{00} + \sum_{i=1}^{\infty} w_i D_{0i} \tag{23}$$

où $w_i(\boldsymbol{x})$ est le coefficient de y_0 dans la somme des termes de $\varphi_i(\boldsymbol{x}, \boldsymbol{y})$ qui contiennent y_0 au premier degré et aucun autre y_i. Mais d'après (18), un tel terme ne peut provenir que du produit $E_{p^i-1}(\boldsymbol{x})\, E_1(\boldsymbol{y})$, et comme $E_1(\boldsymbol{y}) = y_0$, on a $w_i(\boldsymbol{x}) = E_{p^i-1}(\boldsymbol{x})$, donc w_i est isobare et de poids $p^i - 1$. Pour $i = 1$, un calcul immédiat montre que $w_1(\boldsymbol{x}) = -x_0^{p-1}$.

Cela étant, dans le calcul des puissances successives X_{00}^t $(2 \leqq t \leqq p)$ en vue d'obtenir X_{00}^p, seuls importent les coefficients des dérivations D_{0i}, puisque l'on sait d'avance que X_{00}^p est une dérivation. Si $w_i^{(t)}$ est le coefficient de D_{0i}, nous allons voir par récurrence que l'on a $w_1^{(t)} = -(p-1)(p-2)\ldots(p-t+1)\, x_0^{p-t}$, et que pour $i \geqq 2$, $w_i^{(t)}$ est nul ou isobare et de poids $p^i - t$.

En effet, supposons qu'il en soit ainsi; alors, pour $j \geq i$, $D_{0j} w_i^{(t)}$ est nul, et pour $j < i$, $D_{0j} w_i^{(t)}$ est nul, ou isobare et de poids $p^i - t - p^j$. Or, $w_i^{(t+1)}$ est égal à $\sum_{j=0}^{\infty} w_j D_{0j} w_i^{(t)}$, donc chacun des termes est nul ou isobare et de poids $p^i - t - 1$, d'où notre assertion (le calcul de $w_1^{(t)}$ étant immédiat). Pour $t = p$, on a donc bien $X_{00}^p = -(p-1)! X_{01} = X_{01}$, puisque l'on peut remplacer $\boldsymbol{x}$ par $\boldsymbol{e}$ dans l'expression de X_{00}^p, et qu'on obtient ainsi D_{01}.

Supposons maintenant les formules (19) prouvées pour $h < r$, et calculons X_{r0}^p. On a évidemment $X_{k\varepsilon_0}^p = X_{k\varepsilon_1}$ en vertu de l'hypothèse, pour $0 \leq k \leq p^r - 1$, et comme les $Z_{k\varepsilon_0}$ sont combinaisons linéaires des $X_{l\varepsilon_0}$ à coefficients dans $\boldsymbol{F}_p$ ($l < p^r$), $Z_{k\varepsilon_0}^p$ s'obtient en remplaçant chaque X_{hi} par $X_{h,i+1}$ ($h < r$) dans l'expression de $Z_{k\varepsilon_0}$, c'est-à-dire en appliquant à $Z_{k\varepsilon_0}$ l'homomorphisme $\boldsymbol{t}'$; par suite ([5], formule (13)), on a $Z_{k\varepsilon_0}^p = Z_{k\varepsilon_1}$ pour $0 \leq k \leq p^r - 1$. Il en résulte ([5], lemma 2 et formule (6)) que la différence $X_{r0}^p - X_{r1}$ est une *dérivation*. Or, nous allons voir que, pour $r > 0$, il ne figure *aucune dérivation* D_{0j} dans l'expression de X_{ri} à l'aide des D_α (pour tout $i \geq 0$). En effet, le coefficient de D_{0j} serait le coefficient de $y_i^{p^r}$ dans la somme des termes de $\varphi_j(\boldsymbol{x}, \boldsymbol{y})$ qui contiennent $y_i^{p^r}$ au premier degré et aucun autre y_k. D'après (18), un tel terme ne peut provenir que du produit $E_{p^j - p^{i+r}}(\boldsymbol{x}) E_{p^{i+r}}(\boldsymbol{y})$ (si $j > i + r$); mais comme $E_{p^{i+r}}(\boldsymbol{y}) = y_{i+r}$, notre assertion est démontrée. A fortiori, il ne peut figurer aucune dérivation dans un produit X_α qui n'est pas un produit de dérivations X_{0j}, et nous voyons donc que $X_{r0}^p - X_{r1} = 0$, ce qui achève de démontrer les formules (19).

Reste à voir que la loi de composition de H est canonique; il nous suffira de prouver que, si X_α n'est pas un des X_{0j}, dans la formule (6), il n'y a aucune série formelle t_{ε_j} dont le terme constant ne soit pas nul. Supposons en effet ce point établi: alors, pour $\alpha \neq \varepsilon_i$, montrons que le second membre de la formule (8) ne peut contenir aucun X_{0j}, ce qui signifiera que la loi de groupe est canonique. S'il en était autrement, en substituant dans (8) à chaque X_μ son expression tirée de (6), puis en remplaçant $\boldsymbol{x}$ par $\boldsymbol{e}$, on aurait une relation linéaire (à coefficients constants) entre les Z_β, qui devrait donc avoir ses coefficients nuls. Mais comme par hypothèse aucun $Z_{\varepsilon_j} = X_{0j}$ ne peut provenir de l'expression (6) de X_μ si μ n'est pas lui-même de la forme ε_k, les termes en Z_{ε_j} du second membre de (8) ne pourraient provenir que des termes en X_{0j}, et par suite certains auraient dans ce second membre un coefficient non nul, ce qui est absurde si $|\alpha| > 1$.

Pour achever notre démonstration, remarquons que si $h(\alpha) \geq 1$, les séries formelles t_{ε_j} du second membre de (6) sont en fait *nulles*, comme nous l'avons remarqué ci-dessus. On est donc ramené au cas où $X_\alpha = X_{00}^{\alpha_0} X_{01}^{\alpha_1} \dots X_{0k}^{\alpha_k}$, avec $0 \leq \alpha_i \leq p - 1$ pour $0 \leq i \leq k$ et $|\alpha| > 1$. Raisonnons par récurrence sur k, la proposition ayant déjà été établie pour $k = 0$ dans la première partie du raisonnement. Remarquons tout d'abord que pour calculer X_{0i} il faut prendre les coefficients de y_i dans les monômes des $\varphi_j(\boldsymbol{x}, \boldsymbol{y})$ qui contiennent y_i au premier degré et aucun autre y_k, les multiplier par D_{0j} et en faire la somme. La considération des poids des E_k montre que, dans $\varphi_j(\boldsymbol{x}, \boldsymbol{y})$, ces monômes

sont ceux du produit $E_{p^j-p^i}(\boldsymbol{x})\,E_{p^i}(\boldsymbol{y}) = y_i E_{p^j-p^i}(\boldsymbol{x})$; autrement dit, on a

$$X_{0i} = D_{0i} + \sum_{j>i} E_{p^j-p^i}(\boldsymbol{x})\, D_{0j}. \tag{24}$$

Supposons d'abord $\alpha_0 > 0$; on peut compléter les remarques faites au 1° concernant l'expression de $X_{00}^{\alpha_0}$ en remarquant que dans cette expression, seul le terme $D_{00}^{\alpha_0}$ a un coefficient dont le terme constant est $\neq 0$; il est en effet immédiat par récurrence sur α_0 que dans $X_{00}^{\alpha_0}$, le coefficient de D_{01} est celui (en dehors du terme $D_{00}^{\alpha_0}$) dont le poids minimum est le plus petit, et ce poids est $p-\alpha_0 > 0$. Si $X_{01}^{\alpha_1} X_{02}^{\alpha_2} \dots X_{0k}^{\alpha_k} = \sum_\beta v_\beta D_\beta$, il ne pourrait donc y avoir dans X_α des coefficients de dérivations D_{0j} dont le terme constant serait $\neq 0$ que si ces coefficients étaient des séries de la forme $D_{00}^{\alpha_0} v_{\varepsilon_j}$. Or, observons que dans tout monôme d'un polynôme de la forme $E_{ps}(\boldsymbol{x})$ où figure x_0, l'exposant de x_0 est *multiple de* p, car le poids total de E_{ps} et le poids de chacun des x_i d'indice $i \geq 1$ sont multiples de p. Compte tenu des expressions (24), on en conclut aussitôt que les v_β sont des polynômes en $x_0^p, x_1, \dots, x_k$, et par suite $D_{00} v_\beta = 0$ pour tout β. Notre assertion est donc démontrée dans ce cas.

Supposons ensuite que $\alpha_0 = 0$. Pour trouver les coefficients des D_{0j} dans X_α qui ont un terme constant $\neq 0$, on peut évidemment commencer par remplacer x_0 par 0 dans ces coefficients. Or, si on observe que dans les expressions (24) pour $i \geq 1$, D_{00} ne figure pas, on voit que le nouvel opérateur que l'on obtient après la substitution précédente s'obtient aussi en remplaçant *d'abord* x_0 par 0 dans les coefficients de (24), puis en formant le monôme $X_{01}'^{\alpha_1} \dots X_{0k}'^{\alpha_k}$ à partir des opérateurs X_{0i}' ainsi obtenus. Mais il suit de (18) et (21) que l'on a

$$E_{ps}(0, x_1, \dots, x_i, \dots) = E_s(x_1, \dots, x_i, \dots) \tag{25}$$

comme les considérations de poids le montrent aussitôt. On voit par suite que X_{0i}' se déduit de $X_{0,i-1}$ en remplaçant chacun des indices j qui figurent dans $X_{0,i-1}$ par $j+1$ (aussi bien dans les indéterminées que dans les dérivations). L'hypothèse de récurrence prouve par suite que dans X_α aucun des D_{0j} n'a un coefficient dont le terme constant est $\neq 0$, et la proposition 3 est ainsi complètement démontrée.

6. Groupes de type hyperexponentiel.

Nous dirons qu'un groupe *récurrent* G est *de type hyperexponentiel* s'il satisfait aux trois conditions suivantes :

1° G est *abélien*.

2° L'algèbre de Lie $\mathfrak{g}_0$ de G a pour table de multiplication

$$X_{0i}^p = X_{0,i+1} \qquad (i = 0, 1, \dots). \tag{26}$$

On notera que les conditions (26) entraînent, en vertu de l'existence de l'homomorphisme de Frobenius $\boldsymbol{p}'$ [2, n° 13], que l'on a nécessairement, pour tout $h \geq 1$

$$X_{hi}^p = X_{h,i+1} + \sum_\lambda a_\lambda X_\lambda \tag{27}$$

où, dans la sommation, tous les indices λ sont de hauteur $< h$ et n'ont aucune composante d'indice $< i$ qui soit non nulle. Cela étant, la troisième condition imposée aux groupes de type hyperexponentiel est la suivante:

3° Dans la formule (27), pour tout indice λ tel que $a_\lambda \neq 0$, il y a *au moins une* composante de λ d'indice $> i$ qui est *non nulle.*

Remarquons que ces trois conditions sont *invariantes* par tout *isomorphisme récurrent* $\boldsymbol{v}$ de G sur $\overline{G}$, tel que $v_i(\boldsymbol{x}) = x_i + \cdots$ pour tout i. Cela est immédiat pour les deux premières, puisque $\boldsymbol{v}'$ laisse invariante l'algèbre de Lie. En outre, en raison de l'hypothèse sur les premiers termes des v_i, on a $(\boldsymbol{v}')^{-1}(\overline{X}_{hi}) = X_{hi} + \sum_\mu b_\mu X_\mu$, où dans la sommation, $h(\mu) < h$ et les composantes d'indice $< i$ des μ sont toutes nulles. On en déduit que $(\boldsymbol{v}')^{-1}(\overline{X}^p_{hi})$ est une combinaison linéaire $\sum_\nu c_\nu X_\nu$, où dans chaque indice ν de coefficient $c_\nu \neq 0$, toutes les composantes d'indice $< i$ sont nulles et une composante au moins d'indice $> i$ n'est pas nulle; cela signifie encore que cet élément appartient à l'idéal $\mathfrak{m}_i$, noyau du dérivé de l'homomorphisme canonique $\boldsymbol{u}^{(i)}$ de G sur G_i(n° 3). Mais comme $\boldsymbol{v}$ est un isomorphisme récurrent, on a $\overline{\boldsymbol{u}}^{(i)}\boldsymbol{v} = \boldsymbol{v}\,\boldsymbol{u}^{(i)}$, en désignant par $\overline{\boldsymbol{u}}^{(i)}$ l'homomorphisme canonique de $\overline{G}$ sur $\overline{G}_i$. Si $\overline{\mathfrak{m}}_i$ est le noyau du dérivé de $\overline{\boldsymbol{u}}^{(i)}$, on a par suite $\boldsymbol{v}'(\mathfrak{m}_i) = \overline{\mathfrak{m}}_i$, d'où $\overline{X}^p_{hi} = \boldsymbol{v}'\left(\sum_\nu c_\nu X_\nu\right) \in \overline{\mathfrak{m}}_i$, ce que nous voulions démontrer.

La définition que nous venons de donner est justifiée par le théorème suivant:

Théorème 1. *Si G est un groupe de type hyperexponentiel, il existe un isomorphisme récurrent $\boldsymbol{u}$ de G sur le groupe hyperexponentiel H, tel que $u_i(\boldsymbol{x}) = x_i + \cdots$ (termes non écrits de degré > 1) pour tout $i \geq 0$.*

D'après les remarques qui précèdent l'énoncé, on peut supposer que la loi de groupe de G est *canonique* (n° 3, *in fine*). Nous allons démontrer le théorème par «récurrence sur la hauteur» suivant un procédé déjà utilisé dans [5]. Nous supposerons donc que l'hyperalgèbre de G vérifie, non seulement (26), mais aussi les relations (19) *pour tout i et pour $h < r$.* Notons d'abord que, d'après la démonstration du th. d'unicité ([3, th. 2]; voir aussi [4, n° 4]), les algèbres $\mathfrak{s}_r$ ayant même table de multiplication pour G et H, et ces deux groupes ayant des lois canoniques (prop. 3), les coefficients $d_{\alpha\beta\varepsilon_i}$ qui définissent les lois de groupe [formule (11)] sont les mêmes pour G et H, lorsque α et β sont *tous deux de hauteur* $< r$; on en conclut que, pour $h(\alpha) < r$, l'expression de Z_α à l'aide des X_λ a les *mêmes coefficients* dans G et dans H. On a donc, pour le groupe G

$$Z^p_{(\alpha_0, \alpha_1, \ldots)} = Z_{(0, \alpha_0, \alpha_1, \ldots)}$$

pour $h(\alpha) < r$, puisque ces relations ont lieu dans H. D'après [5, lemma 2], on en déduit que $X^p_{ri} - X_{r,i+1}$ est une *dérivation*, donc une combinaison linéaire des X_{0j}; d'ailleurs, en vertu de la condition 3° dans la définition d'un groupe de type hyperexponentiel, on a

$$X^p_{ri} - X_{r,i+1} = \sum_{j \geq i+1} \lambda_{ij} X_{0j} \qquad (i = 0, 1, \ldots). \tag{28}$$

Considérons un homomorphisme φ de G sur un groupe G' défini par

$$x_i' = \varphi_i(x) = x_i + \sum_{s \geq 2} \mu_{is} x_s^{p^s} \qquad (i = 0, 1, \ldots)$$

On déduit de ces formules

$$x_i'(X_{s,i}) = X_{s,i}^{p^i} \quad \text{pour} \quad h < s$$

$$x_i'(X_{s,i}) = X_{s,i}^{p^i} + \sum_{j < s} \mu_{ij} X_{sj}^{p^j}$$

et par suite, en vertu de (28) et de l'hypothèse que G_s

$$(29) \qquad (X_{s,i}^{p^i})^p - X_{s,i+1}^{p^{i+1}} = \sum_{j < s-1} (\lambda_{ij} - \mu_{j,i+1} - \mu_{j-1,i}^p) X_{sj}^{p^j}.$$

On voit aussitôt qu'on peut déterminer les μ_{ij} de façon que les seconds membres de (29) s'annulent pour tout i ou pour tout i sauf un. Les μ_{ij} pour $i \geq 1$ sont déterminés par un système récurrent $(i = 1, 2, \ldots)$. On obtient ainsi un groupe G' de type hyperexponentiel satisfaisant aux relations (29) pour tout i et pour $h \leq r - 1$, la loi de composition n'en pouvant pas rencontrer, mais le procédé employé à fournir un homomorphisme φ_s de $G^{(p)}$ sur G satisfaisant aux mêmes propriétés et ayant même noyau. Le processus de démonstration peut donc se poursuivre et conduit finalement (par le passage à la limite lorsque s tend vers ∞) à un homomorphisme de G sur un groupe H de type hyperexponentiel, ayant même noyau et dont l'hyperalgèbre est définie par (19). En vertu de la prop. 3 et du th. 1 (cas 2), H est identique à H et le théorème est démontré.

Corollaire. Si G est un groupe de type hyperexponentiel, le groupe G_1 est isomorphe à H et le groupe G_s est isomorphe à H_s.

Cela résulte des propriétés générales des homomorphismes (cf. n° 3) et du fait que H_s' est isomorphe à H.

7. Les groupes de Witt.

Le groupe additif de Witt que nous noterons W ou W_∞, est caractérisé (cf. n° 1) de la façon suivante : on considère sur le corps des nombres rationnels, le système récurrent

$$\begin{cases} z_0 = x_0 + y_0 \\ z_0^p + p z_1 = (x_0^p - p x_1) + (y_0^p + p y_1) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ z_0^{p^n} + p z_1^{p^{n-1}} + \cdots + p^n z_n = (x_0^{p^n} + p x_1^{p^{n-1}} + \cdots + p^n x_n) + (y_0^{p^n} + p y_1^{p^{n-1}} + \cdots + p^n y_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \end{cases}$$

On démontre que la solution $z_n = L_n(x_0, y_0, x_1, y_1, \ldots, x_n, y_n)$ $(n = 0, 1, \ldots)$ est telle que L_n soit un polynôme à coefficients entiers, isobare et de poids p^n, la réduction des coefficients mod. p ne faisant rien de la loi de groupe résultant même qui définit W.

De même, le «groupe des Einseinheiten» de Witt (noté V ou V_∞) est défini en considérant le système récurrent (sur le corps des rationnels)

$$(31)\quad \begin{cases} 1+p\,z_0 = (1+p\,x_0)\,(1+p\,y_0) \\ 1+p\,z_0^p+p^2 z_1 = (1+p\,x_0^p+p^2 x_1)\,(1+p\,y_0^p+p^2 y_1) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ 1+p\,z_0^{p^i}+\dots+p^{i+1}z_i = (1+p\,x_0^{p^i}+\dots+p^{i+1}x_i)\,(1+p\,y_0^{p^i}+\dots+p^{i+1}y_i) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \end{cases}$$

Sa solution $z_i = M_i(x_0, y_0, x_1, y_1, \dots, x_i, y_i)$ $(i = 0, 1, \dots)$ est encore telle que M_i soit un polynôme à coefficients entiers, qui ici n'est plus isobare, mais dont le poids minimum est p^i; en réduisant les coefficients mod. p, on obtient la loi de groupe de V.

Les groupes V et W sont des *groupes de type hyperexponentiel*, donc *isomorphes* à H. La vérification des relations (26) peut en effet se faire exactement comme pour le groupe H (n° 5): il suffit d'observer que l'on a

$$(32)\qquad X_{0i} = D_{0i} + \sum_{j \geq i+1} w_{ij} D_{0j}$$

et qu'en raison de la définition des groupes récurrents, le poids minimum de w_{ij} est $\geq p^j - p^i$; en outre, dans $w_{i,i+1}$ figure le monôme $-x_i^{p-1}$, et les autres monômes contiennent tous au moins un x_k d'indice $k<i$. Reste à prouver la troisième condition de la définition d'un groupe de type hyperexponentiel. Mais comme, pour V et W, la série formelle $\varphi_i(\boldsymbol{x}, \boldsymbol{y})$ est de la forme $x_i + y_i + \psi_i(x_0, y_0, \dots, x_{i-1}, y_{i-1})$, X_{hi} est somme de D_{hi} et d'une combinaison linéaire de D_β, où dans chaque β, une au moins des composantes d'indice $> i$ est $\neq 0$. Comme $D_{hi}^p = 0$, X_{hi}^p s'exprime comme combinaison linéaire de D_β où les indices β ont la même propriété, et en remplaçant $\boldsymbol{x}$ par $\boldsymbol{e}$, on en conclut que X_{hi}^p est combinaison linéaire des Z_β correspondant à ces indices, ce qui achève le raisonnement.

Le lecteur observera que nous sommes loin d'avoir utilisé la forme particulière des lois des groupes V et W, et nous lui laissons le soin de formuler des conditions générales permettant d'affirmer qu'un groupe récurrent est du type hyperexponentiel. En fait, on peut obtenir des résultats beaucoup plus précis pour le groupe W, et prouver entre autres que son hyperalgèbre a la même table de multiplication (19) que celle de H; les deux groupes ne sont toutefois pas identiques, mais ont entre eux des relations assez remarquables. Comme ces questions sont un peu étrangères à l'objet principal de ce travail, et comportent des démonstrations assez longues, nous les exposerons dans une publication ultérieure.

8. Le groupe universel.

Nous allons définir encore un autre groupe de type hyperexponentiel, mieux adapté que les précédents aux applications que nous avons en vue:

Proposition 4. *Il existe un groupe U de type hyperexponentiel sur le corps $\boldsymbol{F}_p$ dont l'hyperalgèbre de Lie est définie par les relations* (19), *et dont*

l'opérateur de translation est donné par la formule

$$R_y = \prod_{i=0}^{\infty} \mathrm{Hex}(y_i X_{0i}, y_i^p X_{1i}, \ldots, y_i^{p^h} X_{hi}, \ldots). \tag{33}$$

Nous allons utiliser la prop. 1. Les relations (19) montrent que, du point de vue abstrait, l'hyperalgèbre du groupe dont nous voulons démontrer l'existence est isomorphe à *l'algèbre des polynômes* P (à coefficients dans $\boldsymbol{F}_p$) par rapport aux indéterminées

$$T_0 = X_{00}, \quad T_1 = X_{10}, \ldots, \quad T_h = X_{h0}, \ldots.$$

On a donc

$$\mathrm{Hex}(y_i X_{0i}, y_i^p X_{1i}, \ldots, y_i^{p^h} X_{hi}, \ldots) = \sum_{k=0}^{\infty} y_i^k E_k(T_0^{p^i}, T_1^{p^i}, \ldots, T_h^{p^i}, \ldots).$$

Considérons les polynômes

$$E_\alpha(T_0, \ldots, T_h, \ldots) = \prod_{i=0}^{\infty} E_{\alpha_i}(T_0^{p^i}, \ldots, T_h^{p^i}, \ldots) \tag{34}$$

pour tout indice $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots)$ ($\alpha_i = 0$ à partir d'un certain rang). Montrons d'abord que les E_α forment une *base* de l'algèbre de polynômes P. Soit $k = v_0 + v_1 p + \cdots + v_h p^h$ un entier ≥ 0 développé p-adiquement; si on décompose la série (15) en produit des exponentielles $\exp(u_i)$, on voit aussitôt qu'il y a dans la série (15) un terme de poids k égal à $\frac{u_0^{v_0} u_1^{v_1} \ldots u_h^{v_h}}{v_0!\, v_1! \ldots v_h!}$, et que si on ordonne les monômes en les u_i suivant l'ordre lexicographique *inverse* (le «premier» de deux monômes $u_0^{\alpha_0} \ldots u_h^{\alpha_h} \ldots$, $u_0^{\beta_0} \ldots u_h^{\beta_h} \ldots$ étant $u_0^{\alpha_0} \ldots u_h^{\alpha_h} \ldots$ si et seulement si, pour le *dernier* des indices j tels que $\alpha_j \neq \beta_j$, on a $\alpha_j > \beta_j$), le terme $\frac{u_0^{v_0} \ldots u_h^{v_h}}{v_0! \ldots v_h!}$ est le «premier» des termes de poids k. On en conclut que, lorsqu'on substitue à chaque u_i le polynôme isobare $x_i + G_i(x_0, \ldots, x_{i-1})$, on obtient, comme «premier» terme de poids k, le monôme $\frac{x_0^{v_0} \ldots x_h^{v_h}}{v_0! \ldots v_h!}$; son coefficient n'étant pas multiple de p, c'est aussi ce monôme qui est le «premier» dans $E_k(x_0, \ldots, x_i, \ldots)$. Comme l'ordre lexicographique inverse est compatible avec l'addition dans $\boldsymbol{N}^{(\boldsymbol{N})}$, le «premier» terme de E_α est le produit des «premiers» termes des E_{α_i} dans la formule (34). On en conclut aisément que les E_α sont linéairement indépendants sur $\boldsymbol{F}_p$: il suffit, dans une relation linéaire entre ces polynômes (en nombre fini), de considérer l'ensemble des termes de plus petit poids, dont la somme doit déjà être nulle; on considère ensuite la somme des «premiers» termes des E_α (supposés de même poids), et tout revient à voir que si $\alpha \neq \beta$, les «premiers» termes de E_α et de E_β sont distincts. Mais si on pose $\alpha_i = \lambda_{0i} + \lambda_{1i} p + \cdots + \lambda_{hi} p^h$ pour tous les $\alpha_i \neq 0$, avec $0 \leq \lambda_{ji} \leq p-1$, on peut supposer que h est le même pour tous les α_i (avec éventuellement $\lambda_{hi} = 0$). Si on écrit $m_k = \sum_{i=0}^{\infty} \lambda_{ki} p^i$ ($0 \leq k \leq h$), le «premier» terme de E_α est le monôme $\frac{1}{\alpha!} T_\mu = \frac{1}{\mu!} T_\mu$, où

$$T_\mu = T_0^{m_0} T_1^{m_1} \ldots T_h^{m_h}. \tag{35}$$

Il est clair que les «premiers» termes de E_α et E_β ne peuvent être égaux que si les λ_{ki} sont les mêmes pour $0 \leq k \leq h$, $0 \leq i < +\infty$, ce qui signifie que $\alpha = \beta$.

Reste à montrer que tout monôme (35) est combinaison linéaire des E_α; or, si à partir des m_k on forme les α_i par le procédé inverse de celui décrit ci-dessus, et si on forme $T_\mu - \alpha!\, E_\alpha$, on obtient un polynôme isobare de *même poids* que T_μ, et dont le «premier» terme est «après» celui de E_α. Comme l'ensemble des monômes *de poids donné* est fini, on est ramené, par récurrence, à démontrer notre assertion pour les «derniers» monômes de poids donné (pour l'ordre lexicographique inverse), c'est-à-dire les monômes $T_0^{m_0}$. Or, pour $0 \leq k < p$, on a $E_k(T_0) = T_0^k/k!$; donc $T_0^{m_0}$ est égal à $\prod_{i=0}^{\infty} E_{\lambda_0 i}(T_0^{p^i})$ à un facteur près; ceci achève de démontrer que les E_α forment une base de P.

Ceci étant établi, on peut écrire

$$E_\alpha E_\beta = \sum_\gamma d_{\alpha\beta\gamma} E_\gamma \tag{36}$$

et comme E_k ne contient que $T_0, \ldots, T_{h-1}$ si $k < p^h$, les E_α correspondant aux α de hauteur $< h$ engendrent une sous-algèbre de P, savoir l'algèbre P_h des polynômes en $T_0, \ldots, T_{h-1}$. Pour appliquer la prop. 1, il faut considérer le produit tensoriel $P \otimes P$; ce n'est autre comme on sait (à une isomorphie près), que l'algèbre des polynômes par rapport aux deux séries d'indéterminées $T_0, T_1, \ldots$; $U_0, U_1, \ldots$, et elle a pour base (lorsqu'on fait l'identification précédente) les produits $E_\alpha(T_0, \ldots, T_h, \ldots) E_\beta(U_0, \ldots, U_h, \ldots)$. Or, la formule (17) montre que

$$\sum_{j=0}^{k} E_j(T_0, \ldots, T_h, \ldots) E_{k-j}(U_0, \ldots, U_h, \ldots) = E_k(\Phi_0, \Phi_1, \ldots, \Phi_k, \ldots) \tag{37}$$

où on a posé $\Phi_h = \varphi_h(T_0, \ldots, T_h; U_0, \ldots, U_h)$. De cette formule et de (34), on déduit aussitôt (compte tenu de ce que tous les coefficients sont dans $\boldsymbol{F}_p$)

$$E_\alpha(\Phi_0, \Phi_1, \ldots, \Phi_h, \ldots) = \sum_{0 \leq \beta \leq \alpha} E_\beta(T_0, \ldots, T_h, \ldots) E_{\alpha-\beta}(U_0, \ldots, U_h, \ldots). \tag{38}$$

Mais comme (36) est une identité polynomiale, on peut y substituer Φ_h à T_h pour tout h, et on conclut de (38) que les éléments $\sum_{0 \leq \beta \leq \alpha} E_\beta(T_0, \ldots, T_h, \ldots) E_{\alpha-\beta}(U_0, \ldots, U_h, \ldots)$ forment la base d'une sous-algèbre de $P \otimes P$ (identifiée à l'algèbre des polynômes par rapport aux T_h et aux U_h). Enfin, comme E_α est isobare et de poids $\pi(\alpha)$, il est clair que $d_{\alpha\beta\varepsilon_i} = 0$ si $\pi(\alpha+\beta) < p^i$ ou si une des composantes de α (ou de β), d'indice $j > i$, n'est pas nulle, puisqu'alors $\pi(\alpha) \geq p^j$. Toutes les conditions d'application de la prop. 1 (ou plutôt de son extension aux groupes récurrents, cf. n° 4) sont par suite réalisées, et la prop. 4 est démontrée.

Nous dirons que le groupe récurrent U est le *groupe universel* (terminologie qui sera justifiée au n° 9). Une conjecture naturelle serait que les groupes H et U sont identiques; il n'en est malheureusement rien, comme nous allons

le voir sur un exemple. En raison de la prop. 3 et du fait que les tables de multiplication des hyperalgèbres de H et de U sont les *mêmes*, il suffit de montrer que la loi de groupe de U *n'est pas canonique* en général; autrement dit, il nous suffira de donner un exemple d'un opérateur E_α de l'hyperalgèbre de U, tel que $\alpha \neq \varepsilon_i$, et qui contient un terme en X_{0j} dans son expression à l'aide des X_λ. Il suffit de prendre $p=3$, $\alpha=(6,1,0,0,\ldots)$; on trouve facilement $E_6(T_0, T_1, \ldots) = T_0^6 - T_1^2$, d'où $E_\alpha(T_0, T_1, \ldots) = T_0^9 - T_1^2 T_0^3 = X_{02} - X_{10}^2 X_{01}$.

9. Les «groupes quotients» du groupe universel.

Nous avons démontré dans [5, th. 3] qu'un groupe de Lie formel abélien G de dimension finie sur un corps algébriquement clos est produit direct d'un certain nombre de groupes isomorphes au groupe multiplicatif W_1^* (dont le produit constitue ce que nous avons appelé le *noyau* («core») de G), et d'un *groupe radical*, dont l'hyperalgèbre est une *nilalgèbre* (i.e., tous ses éléments sont nilpotents). Nous nous proposons de montrer comment on peut, à partir du groupe universel U, définir des familles de groupes radicaux, qui en sont des *images homomorphes*, et qui contiennent en particulier tous les groupes radicaux à un paramètre I_r ($r \geq 1$) qui (à l'existence près) avaient été caractérisés dans [5] par leur hyperalgèbre.

Avant de développer la méthode donnant ces groupes, remarquons en passant que le groupe multiplicatif lui-même apparaît comme image homomorphe de U, ou mieux ici de H. En effet, par sa définition même, la série hyperexponentielle $\mathrm{Hex}(x_0, x_1, \ldots)$ constitue un homomorphisme du groupe H sur le groupe multiplicatif W_1^*. Une conséquence immédiate de ce fait est la propriété «différentielle» de cette série, qui la rapproche de l'exponentielle classique: pour *tout opérateur différentiel invariant* Δ du groupe H, on a

$$\Delta\big(\mathrm{Hex}(x_0, x_1, \ldots)\big) = \mathrm{Hex}(x_0, x_1, \ldots). \tag{39}$$

En effet, on a par définition $\Delta R_{\boldsymbol{y}} = R_{\boldsymbol{y}} \Delta$, et ici

$$R_{\boldsymbol{y}}\big(\mathrm{Hex}(x_0, x_1, \ldots)\big) = \mathrm{Hex}(y_0, y_1, \ldots)\,\mathrm{Hex}(x_0, x_1, \ldots).$$

Il suffit d'appliquer Δ aux deux membres, puis, tenant compte de $\Delta R_{\boldsymbol{y}} = R_{\boldsymbol{y}} \Delta$, de faire $\boldsymbol{y} = \boldsymbol{e}$ dans l'équation obtenue, pour avoir (39).

Comme premiers exemples de groupes radicaux déduits du groupe universel, nous allons considérer des groupes dont l'*algèbre de Lie* $\mathfrak{g}_0$ est *indécomposable*, autrement dit [5, n° 11] admet une base telle que $X_{0i}^p = X_{0,i+1}$ pour $0 \leq i < n-1$, $X_{0,n-1}^p = 0$. Soient m et r deux entiers tels que $0 \leq m \leq n-1$, $r \geq 1$. Nous allons montrer qu'*il existe un groupe* $G_{n,m,r}$ *de dimension* n *dont l'hyperalgèbre a pour table de multiplication*

$$\begin{cases} X_{hi}^p = X_{h,i+1} & (h = 0, 1, \ldots;\ 0 \leq i < n-1) \\ X_{h,n-1}^p = 0 \quad \text{pour} \quad h < r \\ X_{r+h,n-1}^p = X_{hm} \quad \text{pour} \quad h = 0, 1, \ldots. \end{cases} \tag{40}$$

Il semble assez commode de représenter cette table de multiplication sous forme d'un *diagramme*, où les X_{hi} sont rangés en un tableau à double entrée suivant la convention habituelle, et où une flèche joignant deux points signifie que l'extrémité de la flèche est la puissance p-ème de l'origine; le diagramme ci-dessous correspond donc au cas $n=4$, $m=1$, $r=3$:

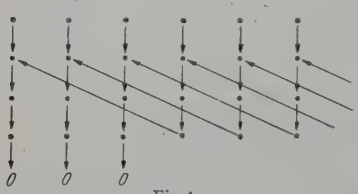

Fig. 1.

Pour démontrer notre assertion, nous remarquerons qu'en supposant le problème résolu, l'hyperalgèbre de $G_{n,m,r}$ sera l'*algèbre quotient* de l'algèbre de polynômes P par l'*idéal* $\mathfrak{a}$ engendré par les polynômes

$$(42) \qquad \begin{cases} T_h^{p^n} & \text{pour } 0 \leq h < r \\ T_h^{p^n} - T_{h-r}^{p^m} & \text{pour } h \geq r. \end{cases}$$

Posons $\overline{P} = P/\mathfrak{a}$, et désignons par $\overline{T}_h$ la classe mod. $\mathfrak{a}$ de T_h. L'algèbre $\overline{P}$ est évidemment engendrée par les éléments $X_{hi} = \overline{T}_h^{p^i}$ pour $0 \leq i \leq n-1$ et $h \geq 0$, et par suite aussi par les éléments

$$(43) \qquad \overline{E}_\alpha(\overline{T}_0, \overline{T}_1, \ldots) = \prod_{i=0}^{n-1} E_{\alpha_i}(\overline{T}_0^{p^i}, \overline{T}_1^{p^i}, \ldots, \overline{T}_h^{p^i}, \ldots)$$

où $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$ ne parcourt cette fois que l'ensemble des indices «de dimension n». En fait, nous allons voir que les $\overline{E}_\alpha$ forment une *base* de $\overline{P}$ (sur le corps $\boldsymbol{F}_p$). Admettons ce point pour un instant; alors on peut écrire

$$(44) \qquad \overline{E}_\alpha \overline{E}_\beta = \sum_\gamma \overline{d}_{\alpha\beta\gamma} \overline{E}_\gamma \qquad (\overline{d}_{\alpha\beta\gamma} \in \boldsymbol{F}_p)$$

et ces identités peuvent être interprétées comme des *congruences*

$$(45) \qquad E_\alpha(T_0, T_1, \ldots)\, E_\beta(T_0, T_1, \ldots) \equiv \sum_\gamma \overline{d}_{\alpha\beta\gamma} E_\gamma(T_0, T_1, \ldots) \qquad (\text{mod. } \mathfrak{a}).$$

Posons alors $\overline{E}^0_\alpha(\overline{T}_0, \overline{T}_1, \ldots; \overline{U}_0, \overline{U}_1, \ldots) = \sum_{0 \leq \beta \leq \alpha} \overline{E}_\beta(\overline{T}_0, \overline{T}_1, \ldots)\, \overline{E}_{\alpha-\beta}(\overline{U}_0, \overline{U}_1, \ldots)$. Il est clair que pour démontrer l'existence du groupe $G_{n,m,r}$, il suffit, en vertu de la prop. 1, de prouver que les $\overline{E}^0_\alpha$ satisfont aux *mêmes* relations (44) que les $\overline{E}_\alpha$. Or, en vertu de (38), on a

$$\overline{E}^0_\alpha(\overline{T}_0, \overline{T}_1, \ldots; \overline{U}_0, \overline{U}_1, \ldots) = E_\alpha(\overline{\Phi}_0, \overline{\Phi}_1, \ldots)$$

où on désigne par $\overline{\Phi}_h$ le polynôme $\varphi_h(\overline{T}_0, \overline{T}_1, \ldots; \overline{U}_0, \overline{U}_1, \ldots)$. Cela étant, si, dans la congruence polynomiale (45), on substitue $\overline{\Phi}_h$ à chaque T_h, on voit que la différence

$$\overline{E}^0_\alpha \overline{E}^0_\beta - \sum_\gamma \overline{d}_{\alpha\beta\gamma} \overline{E}^0_\gamma$$

est combinaison linéaire, à coefficients dans $\overline{P}$, des éléments $\overline{\Phi}_h^{p^n}$ $(0 \leq h < r)$ et $\overline{\Phi}_h^{p^n} - \overline{\Phi}_{h-r}^{p^m}$ (pour $h \geq r$); tout revient à voir que ces différences sont *nulles*.

Mais cela revient à démontrer que l'on a

$$\varphi_h\Big(\overbrace{0,\dots,0}^{r\text{ fois}},T_0^{p^n},\dots,T_h^{p^n},\dots;\overbrace{0,\dots,0}^{r\text{ fois}},U_0^{p^n},\dots,U_h^{p^n},\dots\Big)$$
$$=\varphi_{h-r}(T_0^{p^n},\dots,T_h^{p^n},\dots;U_0^{p^n},\dots,U_h^{p^n}\dots)$$

identité qui résulte de (21) (on a naturellement utilisé le fait que tous les coefficients des polynômes qui interviennent sont dans $\boldsymbol{F}_p$).

Nous sommes donc ramenés à démontrer que les $\overline{E}_\alpha$ forment une base de $\overline{P}$. Pour cela, prouvons d'abord que les formules (40) définissent bien la table de multiplication d'une algèbre associative ayant pour base les monômes X_α. En supposant le problème résolu, on voit immédiatement (notamment sur le diagramme) que l'algèbre cherchée se décomposera en produit tensoriel de r sous-algèbres, dont chacune sera isomorphe à une algèbre engendrée par des éléments X_{hi} satisfaisant aux relations (40) où on a fait $r=1$; tout revient donc à considérer ce cas particulier. Pour cela, envisageons d'abord le cas plus particulier où $r=1$, $m=0$; $\overline{P}$ apparaît alors comme «limite inductive» de la suite d'algèbres

$$A_0=\boldsymbol{F}_p[S_0]/(S_0^p),\ A_1=A_0[S_1]/(S_1^p-s_0),\dots,A_k=A_{k-1}[S_k]/(S_k^p-s_{k-1}),\dots$$

où s_{k-1} désigne la classe de S_{k-1} dans A_{k-1}. Par récurrence sur k, on voit aisément que cette algèbre particulière, que nous désignerons par B, admet pour base les monômes s_α par rapport aux s_k, l'exposant de chacun des s_k étant $\leq p-1$. Notre assertion est donc établie dans ce cas. Pour $r=1$ et $0<m\leq n-1$, formons le «produit tensoriel infini» Q de B et d'une infinité d'algèbres C_h, dont chacune est isomorphe à $\boldsymbol{F}_p[T]/(T^{p^m})$, et a donc une base formée des puissances $t_h^k\ (0\leq k\leq p^m-1)$ d'un élément t_h avec $t_h^{p^m}=0$[2]). L'algèbre Q a donc une base formée des monômes $s_\alpha t_\beta$, où dans t_β l'exposant de chaque t_h est $\leq p^m-1$. Posons alors

$$e_0=s_{n-1}+t_0,\quad e_1=s_{2(n-m)+m-1}+t_1,\dots,\quad e_h=s_{(h+1)(n-m)+m-1}+t_h,\dots\ .$$

Comme on a $e_h^{p^m}=s_{(h+1)(n-m)-1}$, il résulte de ce qui précède que les monômes e_α, où l'exposant de chaque e_h est $\leq p^n-1$, forment une base de l'algèbre Q, et comme $e_h^{p^n}=e_{h-1}^{p^m}$ pour $h\geq 1$, $e_0^{p^n}=0$, Q est isomorphe à l'algèbre $\overline{P}=P/\mathfrak{a}$ lorsque $r=1$.

Pour établir enfin que les $\overline{E}_\alpha$ forment une base de $\overline{P}$, il suffit de remarquer que, d'après (43) et le raisonnement que nous venons de faire, $\overline{E}_\alpha$ s'exprime d'une seule façon comme polynôme par rapport aux $\overline{T}_h$, l'exposant de chacun des $\overline{T}_h$ étant toujours $\leq p^n-1$. En outre, si on ordonne encore ces monômes par rapport aux $\overline{T}_h$ suivant l'ordre lexicographique inverse, le «premier» terme de $\overline{E}_\alpha$ provient du «premier» terme de E_α: cela résulte en effet de ce

[2]) Intuitivement, l'idée de la démonstration consiste à considérer chaque «composante connexe» du diagramme; une telle composante consiste en un graphe linéaire infini (correspondant à l'algèbre B) auquel sont «attachés» aux «noeuds» espacés de $n-m$ en $n-m$, des graphes linéaires de longueur m (correspondant aux algèbres C_h).

que dans le «premier» terme de E_α chaque T_h figure à une puissance $\leq p^n - 1$, puisque $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$ n'a que n composantes, et du fait que les réductions qui s'opèrent dans le passage au quotient, en vertu des relations $\overline{T}_h^{p^n} = \overline{T}_{h-1}^{p^m}$, ne peuvent jamais qu'*augmenter* la place d'un monôme quand on passe de E_α à $\overline{E}_\gamma$. L'indépendance linéaire des $\overline{E}_\alpha$ en résulte alors comme au n° 8, et comme le nombre des $\overline{E}_\alpha$ de hauteur donnée est égal au nombre des monômes X_α de même hauteur, les $\overline{E}_\alpha$ forment une base de $\overline{P}$, ce qui achève de démontrer l'existence des groupes $G_{n,m,r}$.

On notera que, par définition du groupe $G_{n,m,r}$, l'opérateur de translation dans ce groupe est donné explicitement par la formule

$$R_{\boldsymbol{y}} = \prod_{i=0}^{n-1} \mathrm{Hex}(y_i X_{0i}, y_i^p X_{1i}, \ldots, y_i^{p^k} X_{ki}, \ldots) \tag{46}$$

où bien entendu il faut réduire les puissances des X_{hi} à des exposants $< p$ compte tenu des formules (40). En outre, si $\boldsymbol{u}'$ est l'homomorphisme canonique de P sur $\overline{P} = P/\mathfrak{a}$, $\boldsymbol{u}'$ transforme la sous-algèbre $\mathfrak{U}^0$ du produit tensoriel $\mathfrak{U} \otimes \mathfrak{U}$ ($\mathfrak{U}$ hyperalgèbre de U) en la sous-algèbre du produit tensoriel $\mathfrak{G}_{n,m,r} \otimes \mathfrak{G}_{n,m,r}$ ($\mathfrak{G}_{n,m,r}$ hyperalgèbre de $G_{n,m,r}$) engendrée par les $\bar{\Phi}_h$ [en raison de (38)]. Ce qui précède prouve donc (prop. 2) que $\boldsymbol{u}'$ est le dérivé d'un *homomorphisme* $\boldsymbol{u}$ du groupe U sur $G_{n,m,r}$.

Lorsque $n = 1$, on a nécessairement $m = 0$ et les formules (40) montrent que le groupe $G_{1,0,r}$ est *isomorphe au groupe* I_r défini dans [5, n° 8] et dont l'existence se trouve ainsi établie.

On notera d'autre part que $G_{n,m,r}$ et $G_{n',m',r'}$ *ne peuvent être isomorphes que si* $n = n'$, $m = m'$ *et* $r = r'$. En effet ,l'*indice de nilpotence* de l'algèbre $\mathfrak{s}_h$ des semi-dérivations spéciales invariantes de hauteur $< h$ (plus petit exposant e tel que $Y^e = 0$ pour tout élément Y de l'algèbre) est égal à p^n pour $h \leq r$, à p^{2n-m} pour $h = r + 1$, et par suite les nombres m et r sont déterminés intrinsèquement par le groupe $G_{n,m,r}$.

10.

Comme deuxième exemple de groupes obtenus à partir du groupe U, considérons des groupes de dimension 2, dont l'algèbre de Lie $\mathfrak{g}_0$ est «complètement triviale», c'est-à-dire telle que la puissance p-ème de tout élément de $\mathfrak{g}_0$ soit nulle. On a évidemment, parmi ces groupes, tous les produits directs $I_r \times I_s$ ($r > 0$, $s > 0$), $I_r \times W_1$ et $W_1 \times W_1$. En outre, nous allons voir que l'on a encore des groupes correspondant aux trois types de diagrammes suivants:

I. *Type «strie»*:

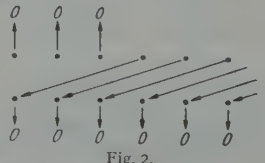

Fig. 2.

En formules, cela correspond aux relations

$$(47)\qquad \begin{cases} X_{h0}^{p} = 0 \quad \text{pour} \quad h < r, \qquad X_{h-r,0}^{p} = X_{h1} \quad \text{pour} \quad h \geq 0 \\ X_{h1}^{p} = 0 \quad \text{pour tout} \quad h \geq 0. \end{cases}$$

II. *Type «arbre»*:

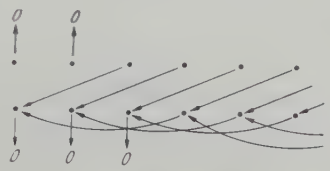

Fig. 3.

ou, en formules:

$$(48)\qquad \begin{cases} X_{h0}^{p} = 0 \quad \text{pour} \quad h < r, \qquad X_{h-r,0}^{p} = X_{h1} \quad \text{pour} \quad h \geq 0 \\ X_{h1}^{p} = 0 \quad \text{pour} \quad h < s, \qquad X_{h+s,1}^{p} = X_{h1} \quad \text{pour} \quad h \geq 0. \end{cases}$$

III. *Type «tresse»*:

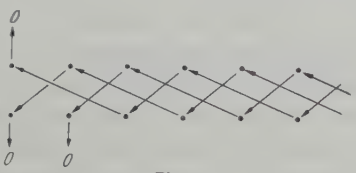

Fig. 4.

ou, en formules:

$$(49)\qquad \begin{cases} X_{h0}^{p} = 0 \quad \text{pour} \quad h < r, \qquad X_{h-r,0}^{p} = X_{h1} \quad \text{pour} \quad h \geq 0 \\ X_{h1}^{p} = 0 \quad \text{pour} \quad h < s, \qquad X_{h+s,1}^{p} = X_{h0} \quad \text{pour} \quad h \geq 0. \end{cases}$$

L'existence des groupes de ces types se démontre en partant cette fois du groupe produit $U \times U$, dont l'hyperalgèbre est $P \otimes P$, et peut par suite être considérée comme algèbre des polynômes à 2 séries de variables $T_0, T_1, \ldots, T_0', T_1', \ldots$, ayant pour base les polynômes $E_\alpha(T_0, T_1, \ldots) E_\beta(T_0', T_1', \ldots)$. Si on veut par exemple démontrer l'existence des groupes du type «tresse», on considérera l'idéal $\mathfrak{b}$ de $P \otimes P$ engendré par les polynômes

$$T_h^p \quad \text{pour} \quad h < r, \qquad T_h^p - T_{h-r}' \quad \text{pour} \quad h \geq r$$
$$T_h'^p \quad \text{pour} \quad h < s, \qquad T_h'^p - T_{h-s} \quad \text{pour} \quad h \geq s.$$

Si $\overline{T}_h, \overline{T}_h'$ sont les classes de T_h, T_h' mod. $\mathfrak{b}$, on montre d'abord que les éléments $\overline{E}_\lambda(\overline{T}_0, \overline{T}_1, \ldots; \overline{T}_0', \overline{T}_1', \ldots) = E_k(\overline{T}_0, \overline{T}_1, \ldots) E_l(\overline{T}_0', \overline{T}_1', \ldots)$, où $\lambda = (k, l)$ parcourt tous les couples d'entiers ≥ 0, forment une *base* de $(P \otimes P)/\mathfrak{b}$; la méthode est analogue à celle du n° 9, et consiste à «dérouler» la «tresse» en ses composantes connexes, autrement dit, à prouver que $(P \otimes P)/\mathfrak{b}$ est isomorphe au produit tensoriel de $r + s$ algèbres identiques à l'algèbre B définie au n° 9. Ceci montre que les monômes X_α par rapport aux $X_{h0} = \overline{T}_h$ et $X_{h1} = \overline{T}_h'$, ayant tous leurs exposants $< p$, forment une base d'une algèbre isomorphe à $(P \otimes P)/\mathfrak{b}$, et vérifiant (49). Pour voir ensuite que les $\overline{E}_\lambda$ forment aussi une base de cette algèbre, il suffit d'ordonner les $\overline{T}_h$ et $\overline{T}_h'$ suivant l'ordre: $\overline{T}_0, \overline{T}_0', \overline{T}_1, \overline{T}_1', \ldots, \overline{T}_h, \overline{T}_h', \ldots$, puis les monômes X_α suivant l'ordre lexicographique inverse (ce n'est autre que l'ordre lexicographique inverse «par

étages» décrit au n° 3); le raisonnement se poursuit alors comme au n° 9. Reste enfin à montrer que la prop. 1 est applicable; on voit comme au n° 9 que cela revient à prouver que, si on pose

$$\overline{\Phi}_h = \varphi_h(\overline{T}_0, \overline{T}_1, \ldots; \overline{U}_0, \overline{U}_1, \ldots), \quad \overline{\Phi}'_h = \varphi_h(\overline{T}'_0, \overline{T}'_1, \ldots; \overline{U}'_0, \overline{U}'_1, \ldots)$$

on a $\overline{\Phi}_h^p = 0$ pour $h < r$, $\overline{\Phi}_h^p = \overline{\Phi}'_{h-r}$ pour $h \geq r$, $\overline{\Phi}_h'^p = 0$ pour $h < s$, $\overline{\Phi}_h'^p = \overline{\Phi}_{h-s}$ pour $h \geq s$; cela se fait encore en utilisant la formule (21).

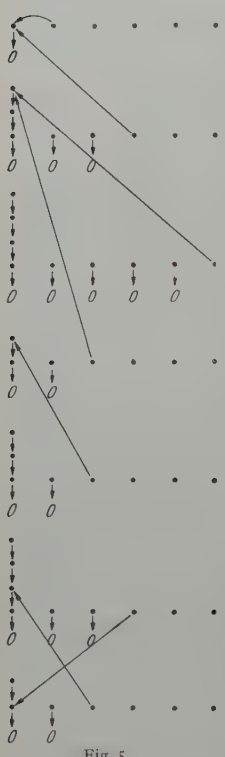

Fig. 5.

Les raisonnements sont très analogues pour les deux autres types. Les considérations d'indice de nilpotence montrent en outre que les groupes de type «tresse dissymétrique» ($r < s$) ne sont pas isomorphes deux à deux ni à aucun groupe des autres types; par contre, on peut prouver que le type «tresse symétrique» ($r = s$) est isomorphe au produit $I_r \times I_r$ (tout au moins sur un corps de base algébriquement clos). De même, on montre que les groupes de type «strie» sont deux à deux non isomorphes et non isomorphes aux autres types; enfin, les mêmes considérations s'appliquent au type «arbre» lorsque $r < s$; j'ignore par contre si pour $r \geq s$, un groupe du type «arbre» ne peut être isomorphe à $W_1 \times I_s$.

11.

Les considérations des deux n^{os} précédents admettent évidemment une généralisation étendue. On forme ainsi des groupes dont l'algèbre de Lie $\mathfrak{g}_0$ est somme directe d'un nombre arbitraire d'algèbres *indécomposables* (on sait [5, n° 11] que toute algèbre de Lie «sans noyau» est de cette forme). A chacune de ces algèbres de Lie indécomposables, nous attacherons un «diagramme simplifié» ne comportant qu'une ligne horizontale et une ligne de flèches verticales; puis nous relierons ces diagrammes par des flèches obliques (éventuellement horizontales), suivant les pricipes suivants: 1° de chaque ligne horizontale, il ne part qu'une seule flèche oblique au plus; 2° la flèche oblique aboutit en un point d'une ligne verticale. On complète ensuite le diagramme en opérant sur chaque flèche une infinité de translations d'une unité vers la droite (non représentées sur la figure qui suit) et en ajoutant à chaque ligne horizontale des flèches aboutissant à 0, avant le début de la première flèche oblique:

L'interprétation des diagrammes est la même que ci-dessus: s'il y a k diagrammes, dont les nombres de lignes sont n_i ($1 \leq i \leq k$), les éléments X_{hij} ($1 \leq i \leq k$, $1 \leq j \leq n_i$) des h-èmes colonnes de ces diagrammes sont les semi-dérivations invariantes de hauteur h telles que $X_{hij}(\boldsymbol{e}) = D_{hij}(\boldsymbol{e})$, et une flèche joint une semi-dérivation à sa puissance p-ème.

Il est naturel de conjecturer que l'on obtient ainsi *tous* les groupes abéliens radicaux (à des isomorphismes près); mais je ne sais pour le moment confirmer ou infirmer cette conjecture[3]). Au cas où elle serait exacte, une étude détaillée des diagrammes précédents et des isomorphismes possibles entre les groupes qui leur correspondent, serait naturellement indispensable.

Bibliographie.

[*1*] DIEUDONNÉ, J.: Sur les groupes de Lie algébriques sur un corps de caractéristique $p > 0$. Rend. Circ. Mat. Palermo (2) **1**, 380—402 (1952). — [*2*] DIEUDONNÉ, J.: Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$. Comm. Math. Helv. **28**, 87—118 (1954). — [*3*] DIEUDONNÉ, J.: Sur la notion de variables canoniques. Anais Acad. brasil. Ciencias **27** (1955). — [*4*] DIEUDONNÉ, J.: Le calcul différentiel dans les corps de caractéristique $p > 0$. Proc. Int. Congress of Math., Amsterdam **1** (1955). — [*5*] DIEUDONNÉ, J.: Lie groups and Lie hyperalgebras over a field of characteristic $p > 0$ (II). Amer. J. Math. **77**, 218—244 (1955). — [*6*] WITT, E.: Zyklische Körper und Algebren der Charakteristik p vom Grad p^n. J. reine u. angew. Math. **176**, 126—140 (1937).

[3]) *Note ajoutée pendant la correction des épreuves.* Depuis que cet article a été écrit, j'ai obtenu un procédé général de description des groupes abéliens qui comprend comme cas particuliers ceux envisagés ici; voir un article à paraître dans l'*American Journal of Mathematics.*

Dept. of Math., Northwestern University, Evanston, Ill. (USA).

(Eingegangen am 8. November 1954.)

NEUMANN, B. H.
Math. Zeitschr. Bd. 63, S. 76—96 (1955)

Groups with finite classes of conjugate subgroups.

(In memoriam ISSAI SCHUR.)

By
B. H. NEUMANN.

§ 1. Introduction.

In the present paper we shall be concerned with the interrelations of various classes of groups defined by certain properties, and to avoid repetitive circumlocutions we make the following convention. The properties of groups we are mainly interested in will be denoted by short sequences of capitals. If PR is a property of groups, the groups that have the property PR will be called PR-groups, and the class of all PR-groups will be denoted by $[PR]$. Thus, for instance, we denote by FG the property to be "finitely generated" (more accurately: to be capable of being finitely generated; for such groups can, of course, be given in terms of an infinite set of generators). Then $[FG]$ is the class of all FG-groups, that is the class of all finitely generated groups.

The most important properties of groups G to be considered are the following[1]).

FC: The classes of conjugate elements of G are finite.

FD: The derived group G' is finite.

FIZ: The centre has finite index in G.

X: The classes of conjugate subgroups of G are finite.

Y: Every subgroup H of G has finite index in some normal subgroup of G (and therefore, in particular, in its normal closure H^G in G).

FC-groups have recently received attention from several authors[2]). FD is a more restrictive property than FC, that is to say, all FD-groups are FC-groups; FIZ is still more restrictive, that is to say, all FIZ-groups are FD-groups. For finitely generated groups the three properties are co-extensive. We can formulate these (known) relations thus:

$$[FC] \supset [FD] \supset [FIZ]; \tag{1.1}$$

$$[FG] \cap [FC] = [FG] \cap [FD] = [FG] \cap [FIZ]. \tag{1.2}$$

The relations of X and Y to each other and to the other properties are far from obvious. One might try to be guided by the extreme case that arises if "finite" is replaced by "of cardinal 1". Then FC-groups, FD-groups and FIZ-groups all simply become abelian groups, whereas X-groups and Y-

[1]) Further properties, namely BFC, BX, BY, are defined at the end of this section. The classes $[H]$—$[K]$ introduced in §§ 5—10 are of ephemeral significance only.

[2]) Cf. BAER [1], NEUMANN [9], HAIMO [5], [6], ERDÖS [4] and the literature there quoted; also NEUMANN [10].

groups become the groups in which every subgroup is normal. These groups, first studied by DEDEKIND [3], form a somewhat larger class than that of abelian groups, comprising the abelian and the hamiltonian groups. It is then not unnatural to expect that analogously $[FC]$ is contained in $[X]$ and in $[Y]$. This is, however, not the case; on the contrary, we shall show that the X-groups are precisely the FIZ-groups, and the Y-groups are precisely the FD-groups. Thus the interrelations of these classes of groups can be completely—and simply—described. In particular it follows that every X-group is a Y-group; I know no obvious reason why this should be so, nor do I know any direct proof.

A corollary of our results is that in an X-group the classes of conjugate subgroups are not only finite, but boundedly (or uniformly) finite; and that in a Y-group every subgroup has not only finite, but boundedly finite index in some normal subgroup. Let us denote the properties obtained from FC, X, and Y by strengthening "finite" to "boundedly finite" by BFC, BX, BY respectively. Then $[BFC]$ is a proper subclass of $[FC]$; in fact it has been shown [*10*, Theorem 3.1] that $[BFC] = [FD]$, and we shall here rederive this result by a quite different method. On the other hand we show that $[BX] = [X]$ and $[BY] = [Y]$. The results of this paper can then be summarized in the form

$$(1.3) \qquad [FC] > [BFC] = [FD] = [Y] = [BY] > [FIZ] = [X] = [BX].$$

§ 2. Preliminaries.

We use the following notation. If G is a group, then G' is its derived group, Z its centre. If S is a subset of G, then $|S|$ denotes its cardinal; $\{S\}$ is the subgroup it generates; $N(S)$ is its normalizer and $C(S)$ its centralizer in G. If S consists of $g_1, g_2, \ldots, g_n$, then we also write $\{g_1, g_2, \ldots, g_n\}$ for $\{S\}$, $N(g_1, g_2, \ldots, g_n)$ for $N(S)$, and $C(g_1, g_2, \ldots, g_n)$ for $C(S)$. We note that

$$C(\{S\}) = C(S) = \bigcap_{s \in S} C(s),$$

and if S is a generating set of G, then

$$C(S) = C(G) = Z.$$

In general, however, $N(\{S\})$ is not the same as $N(S)$. For a single element, $N(g) = C(g)$.

If H is a subgroup of G, then $|G:H|$ denotes its index in G, and H^G its normal closure in G, that is the least normal subgroup of G containing H. The unit element of all the groups that occur will be denoted by 1.

The class of conjugates of a subset S of G consists of the subsets $g^{-1}Sg$, where g ranges over G. Its cardinal is the index of the normalizer, $|G:N(S)|$. In particular the cardinal of the class of conjugates of an element g is $|G:C(g)|$.

The following known facts on FC-groups will be required [*9*, § 5].

Lemma 2.1. *If G is an FC-group then the elements of finite order form a subgroup P of G; if G is finitely generated then P is finite; P contains G', and*

thus all commutators in G have finite order. Every finite set of elements of finite order generates a finite subgroup, and is even contained in a finite normal subgroup. If S is a finite subset of G then $|G:C(S)|$ is finite.

Lemma 2.2. *Each of the classes $[FC]$, $[BFC]$, $[FD]$, $[FIZ]$, $[X]$, $[BX]$, $[Y]$, $[BY]$ is closed under the operations of forming subgroups and homomorphic images.*

The proof is obvious and omitted.

Lemma 2.3. *Every FIZ-group is an X-group and even a BX-group; every FD-group is a Y-group and even a BY-group; that is to say,*

$$[FIZ] \subseteq [BX] \subseteq [X]; \quad [FD] \subseteq [BY] \subseteq [Y].$$

Proof. The centre Z is contained in every centralizer and *a fortiori* in every normalizer; thus if $G \in [FIZ]$, that is if $|G:Z|$ is finite, then so is $|G:N(H)|$ for every subgroup H, and $G \in [X]$; in fact $|G:Z|$ is a bound for the number of conjugates of every subgroup, and thus $G \in [BX]$. Hence $[FIZ] \subseteq [BX]$. Next, if H is a subgroup of G, then HG' is a normal subgroup of G. Also $|HG':H| \leq |G'|$; if $G \in [FD]$, that is if $|G'|$ is finite, then so is $|HG':H|$ for every H, and $G \in [Y]$; in fact $|G'|$ is a bound for the index of H in one normal subgroup of G, and thus $G \in [BY]$. Thus $[FD] \subseteq [BY]$, and the lemma follows.

§ 3. X-groups and Y-groups are FC-groups.

Lemma 3.1. *If g is an element of a group G then $|N(\{g\}):C(g)|$ is finite.*

Proof. The cyclic group $\{g\}$ is normal in $N(\{g\})$. Each conjugate of g in $N(\{g\})$ then lies in $\{g\}$ and generates $\{g\}$; but a cyclic group has only a finite number of distinct single generators [namely $\varphi(n)$ if $\{g\}$ is of finite order n, and 2 if $\{g\}$ is of infinite order]. Hence g has only a finite number of distinct conjugates in $N(\{g\})$. Finally we observe that the centralizer of g in $N(\{g\})$ is the same as that in G; and the lemma follows.

Lemma 3.2. *Every X-group is an FC-group.*

Proof. If $G \in [X]$ and $g \in G$, then $|G:N(\{g\})|$ is finite because $\{g\}$ has only finitely many distinct conjugates, and $|N(\{g\}):C(g)|$ is finite by the preceding lemma. Hence

$$|G:C(g)| = |G:N(\{g\})| \cdot |N(\{g\}):C(g)|$$

is finite, and g has only a finite number of distinct conjugates. This is true for every $g \in G$; hence $G \in [FC]$, and the lemma follows.

Corollary 3.3. $$[FG] \cap [X] = [FG] \cap [FC].$$

This follows from (1.2) and Lemmas 2.3 and 3.2.

Lemma 3.4[3]. *A finitely generated group H contains only finitely many subgroups K of given finite index $|H:K| = k$.*

[3]) This lemma is known; cf. MARSHALL HALL [7], who determines the precise number when H is free, and thereby provides a bound for arbitrary H. Cf. also BAER [2, p. 331].

Proof. Choose a fixed finite set of generators of H. Every subgroup K of index k determines an equivalence relation on H, namely the right congruence modulo K. This equivalence is completely determined when it is specified for all words of length not exceeding some fixed bound l, and every such word of length l is equivalent to a shorter word. If $h_1 h_2 \dots h_k$ is a word of length k (the h_i being generators or their inverses), then among the $k+1$ elements

$$1, h_1, h_1 h_2, \dots, h_1 h_2 \dots h_k$$

there must be two that are right congruent modulo K, and it easily follows that then $h_1 h_2 \dots h_k$ is right congruent to a shorter word modulo K. Hence we can put $l = k$. Conversely, the right congruence determines K, because K simply consists of all elements right congruent to 1. As there are only a finite number of elements of length not exceeding k, and only a finite number of equivalence relations on a finite set, the number of subgroups K is finite, and the lemma follows.

Lemma 3.5. *If a finitely generated subgroup K of G has finite index in a normal subgroup H of G, then K has only finitely many distinct conjugates in G.*

Proof[4]). We first note that H is itself finitely generated, namely by a finite generating set of K and one representative out of each right coset of K in H. Now every conjugate of K in G is a subgroup of H and has the same index $k = |H:K|$ in H; by the preceding lemma there are only a finite number of such subgroups, and the lemma follows.

Lemma 3.6. *If a cyclic group $\{g\}$ has finite index in its normal closure $\{g\}^G$, then its generator g has only finitely many distinct conjugates in G.*

Proof. By the preceding lemma, $|G:N(\{g\})|$ is finite, and then $|G:C(g)|$ is finite by Lemma 3.1.

Lemma 3.7. *Every Y-group is an FC-group.*

This is an immediate consequence of the definition of $[Y]$ together with Lemma 3.6.

Corollary 3.8. $$[FG] \cap [Y] = [FG] \cap [FC].$$

This follows from (1.2) and Lemmas 2.3 and 3.7.

§ 4. Some properties of FC-groups.

In the present section we show that if an FC-group G contains a sufficiently "large" subgroup in $[FD]$ or in $[FIZ]$, then G is itself an FD-group or FIZ-group, respectively. We also demonstrate a method of replacing a group by a factor group without disturbing any of the properties we are interested in. All these results will be found useful at a later stage.

Lemma 4.1. *Let $G \in [FC]$ possess a subgroup $H \in [FD]$ and a finite subset S such that H and S together generate G. Then $G \in [FD]$.*

[4]) I am indebted to Graham Higman for a simplification of the proof of this lemma.

Proof. The derived group G' is the normal closure of the group generated by the commutators $[h, h']$, $[h, s]$, and $[s, s']$, where h, h' range over H and s, s' over S. There are only finitely many $[h, h']$ because H' is finite; there are only finitely many commutators $[g, s]$ for any one s and for g ranging over G, because such a commutator is the quotient of a conjugate of s—capable of only finitely many different values—and s; thus there are altogether only finitely many commutators of the form $[g, s]$ with s ranging over S. It follows that G' is the normal closure of a finite set of elements of finite order and hence of finite order itself (cf. Lemma 2.1). Thus $G \in [FD]$ and the lemma follows.

Corollary 4.2. *If G is an FC-group but not an FD-group, then no subgroup of finite index in G is an FD-group.*

Lemma 4.3. *Let $G \in [FC]$ possess a subgroup $H \in [FIZ]$ and a finite subset S such that H and S together generate G. Then $G \in [FIZ]$.*

Proof. Denote by K the intersection of H and the centre Z of G:

$$K = H \cap Z.$$

As $G = \{H, S\}$, then

$$Z = C(H) \cap C(S),$$

and therefore

$$K = H \cap C(H) \cap C(S) = (H \cap C(H)) \cap (H \cap C(S)).$$

Now $H \cap C(H)$ is the centre of H, and this has finite index in H by the assumption that $H \in [FIZ]$. Also

$$|H : H \cap C(S)| \leqq |G : C(S)|,$$

and this is finite by Lemma 2.1, as S is a finite set. Hence K, as the intersection of two subgroups of finite index in H, has finite index in H. Denote by T a subset of H which together with K generates H. Then $G = \{K, T, S\}$ and

$$Z = C(K) \cap C(T) \cap C(S).$$

Here $C(S)$ and $C(T)$ have finite index in G because S and T are finite sets (cf. Lemma 2.1); also $C(K) = G$, as $K \leqq Z$. It follows that $|G : Z|$ is finite, and $G \in [FIZ]$, as asserted.

Corollary 4.4. *If G is an FC-group but not an FIZ-group, then no subgroup of finite index in G is an FIZ-group.*

Lemma 4.5. *In an FC-group every element has finite order modulo the centre.*

Proof. Let $g \in G \in [FC]$; then $|G : C(g)|$ is finite. Let $g_1, g_2, \ldots, g_n$ be a complete set of coset representatives of $C(g)$ in G. Then also $|G : C(g_1, g_2, \ldots, g_n)|$ is finite. Hence there is a positive integer m such that $g^m \in C(g_1, g_2, \ldots, g_n)$. As g^m commutes with all the elements that commute with g, that is with all the elements of $C(g)$, and also with $g_1, g_2, \ldots, g_n$, then $g^m \in Z$, and the lemma follows.

Definition. Let φ be a homomorphic mapping of a group G onto a group $G^0 = G\varphi$, and let K be the kernel of φ. If $K \cap G' = \{1\}$ and hence $K \leqq Z$, then we call φ an *isoclinic* homomorphism.

If φ is an isoclinic homomorphism, then the centre Z^0 of G^0 is $Z\varphi$; for if $g \in G$ is not in the centre Z, then there is an element $h \in G$ such that $[g, h] \neq 1$. But then $[g, h] \notin K$, because $K \cap G' = \{1\}$, and therefore $[g\varphi, h\varphi] = [g, h]\varphi \neq 1$. Thus $g\varphi$ is outside the centre of $G\varphi$. It follows that the factor groups of the centres of G and G^0 are isomorphic; for

$$G/Z \cong G\varphi/Z\varphi = G^0/Z^0$$

as Z contains the kernel of φ. In fact φ itself induces an isomorphism of these factor groups; and this has clearly the property of inducing an isomorphism of the derived groups G' and $G^{0\prime}$ also, because $K \cap G' = \{1\}$ entails

$$G^{0\prime} = (G\varphi)' = G'\varphi.$$

All this together shows that G and G^0 are isoclinic in the sense of PHILIP HALL [8]. The isoclinic homomorphism is not the most general isoclinism between groups, but it is sufficient for our purposes.

Lemma 4.6. *Let φ be an isoclinic homomorphism of a group G onto $G^0 = G\varphi$. Then if $G^0 \in [PR]$ where $[PR]$ is one of the classes $[FC]$, $[FD]$, $[FIZ]$, then also $G \in [PR]$.*

Proof. As the derived groups of G^0 and of G are isomorphic, $G^0 \in [FD]$ implies $G \in [FD]$. As the factor groups of the centres of G^0 and G are isomorphic, $G^0 \in [FIZ]$ implies $G \in [FIZ]$. We next show that $g \in G$ has precisely as many distinct conjugates as $g^0 = g\varphi$ has in G^0. Let $h \notin C(g)$. Then $[h, g] \neq 1$, and therefore $[h, g] \notin K$; thus $[h\varphi, g\varphi] = [h, g]\varphi \neq 1$, and $h^0 = h\varphi$ is not in the centralizer $C^0(g^0)$ of g^0 in G^0. Thus nothing outside $C(g)\varphi$ can belong to $C^0(g^0)$; but clearly $C(g)\varphi \subseteq C^0(g^0)$. As, moreover, $C(g)$ contains Z, and thus also K, then

$$|G : C(g)| = |G\varphi : C(g)\varphi| = |G^0 : C^0(g^0)|.$$

It follows that the class of conjugates of an element $g \in G$ has the same cardinal as that of g^0 in G^0; in particular, if $G^0 \in [FC]$, then $G \in [FC]$. This completes the proof of the lemma. It may be noted that the lemma provides a partial converse of Lemma 2.2.

Lemma 4.7. *To every group G there is an isoclinic homomorphism onto a group G^0 in which every element $z^0 \neq 1$ of the centre Z^0 has a power $z^{0m} \neq 1$ in the derived group $G^{0\prime}$.*

Proof. Among the kernels of isoclinic homomorphisms of G, that is among the subgroups of Z which have trivial intersection with G', we choose a maximal one and denote it by K. The existence of such a maximal subgroup follows by a familiar application of ZORN's Lemma. We put $G^0 = G/K$; the canonic homomorphism of G onto G^0 is then isoclinic. If Z^0 is the centre of G^0, then we know (from a remark following the definition of an isoclinic homomorphism) that $Z^0 = Z/K$. Now every non-trivial subgroup of Z^0 must intersect the derived group $G^{0\prime}$ non-trivially; this is an easy consequence of the maximality of K. Thus if $z^0 \in Z^0$, $z^0 \neq 1$, then there is an integer m such that $1 \neq z^{0m} \in G^{0\prime}$, and the lemma follows.

Corollary 4.8. *If $G \in [FC]$, then G^0—chosen according to Lemma 4.7—is periodic, that is to say, all its elements have finite order; and G^0 is even locally finite, that is to say, every finite subset is contained in a finite subgroup.*

Proof. As G^0 is a homomorphic map of G, then $G^0 \in [FC]$. By Lemma 4.5 every element $g^0 \in G^0$ has finite order modulo Z^0. If $g^{0^n} = z^0 \in Z^0$, then g^0 has certainly finite order if $z^0 = 1$; if $z^0 \neq 1$ then $g^{0^{n\,m}} = z^{0^m} = c^0 \in G^{0\prime}$ by Lemma 4.7, and c^0 has finite order by Lemma 2.1. Hence g^0 has finite order, and by a further application of Lemma 2.1 the corollary follows.

§ 5. The class *[H]*.

In this and the following four sections we introduce and study some rather restricted classes of FC-groups which will later be shown to be typical of certain subgroups, or certain factor groups of subgroups, that occur among the groups we are concerned with.

We denote by $[H]$ the class of groups G generated by (denumerably) infinitely many elements

$$a_1, b_1, c_1, a_2, b_2, c_2, \ldots$$

subject to the relations

$$[a_i, b_i] = c_i \neq 1, \qquad [a_i, a_j] = [a_i, b_j] = [b_i, b_j] = 1, \tag{5.1}$$

where $i, j = 1, 2, \ldots$ and $i \neq j$; moreover each subgroup

$$N_i = \{a_i, b_i\}\ (= \{a_i, b_i, c_i\})$$

is to be an FC-group.

The relations (5.1) are not to be understood to form a system of defining relations of G: there will in general be further relations, not consequences of (5.1); but they must not imply $c_i = 1$ for any i. The sequence $a_1, b_1, c_1, a_2, b_2, c_2, \ldots$ will be called a system of H-generators of G.

The following properties of an H-group G are readily verified from the definition. The subgroups N_i are normal in G, and they jointly generate G. The centralizer $C(N_i)$ is generated by all the subgroups N_j with $j \neq i$ together with the centre of N_i. If an element of N_i is also generated by the N_j with $j \neq i$, then it belongs to the centre Z of G; for *qua* element of N_i it is permutable with the elements of all N_j other than N_i, and *qua* product of elements out of the N_j other than N_i it is also permutable with the elements of N_i. The centre Z is generated by the centres of all the N_i. We can consider G as the restricted direct product of the N_i with certain subgroups of their centres amalgamated: but this interpretation will not be required in what follows. Putting

$$F_i = N_i Z / Z,$$

one easily sees that G/Z is the restricted direct product of the groups F_i. Each F_i is a finite group, because N_i, being a finitely generated FC-group, is a FIZ-group [cf. (1.2)], and thus also $|N_i Z : Z|$ is finite. On the other hand G is not a FIZ-group; for none of $a_1, a_2, \ldots$ belong to the centre, and they

are in fact all incongruent modulo Z: if $i \neq j$, then

$$b_i^{-1} a_i a_j^{-1} b_i = a_i a_j^{-1} c_i \neq a_i a_j^{-1},$$

and thus $Z a_i \neq Z a_j$. However, G is an FC-group; for every element of G is of the form

$$g = n_1 n_2 \dots n_q, \qquad n_i \in N_i,$$

q being taken sufficiently large; the conjugates of g then are of the form

$$g' = n_1' n_2' \dots n_q'$$

where n_i' is conjugate to n_i in N_i, and thus capable of finitely many different values only. Hence the class of conjugates of g is finite, and G is an FC-group. We have thus shown

$$[H] \leqq [FC] - [FIZ]. \tag{5.2}$$

§ 6. The class [*I*].

Next we introduce a subclass of $[H]$: we denote by $[I]$ the class of H-groups G that satisfy, in addition to (5.1), the relations

$$[a_i, c_i] = [b_i, c_i] = 1, \qquad c_i^{p^{\varrho(i)}} = 1, \tag{6.1}$$

where $i = 1, 2, \dots$ and p is a prime number. Again there may be further relations, but they must not imply $c_i = 1$ for any i; and we may take $p^{\varrho(i)}$ to be precisely the order of c_i.

The following facts are easy consequences of the definition. In an I-group G the subgroup

$$C = \{c_1, c_2, \dots\}$$

is the derived group; it is contained in the centre, in other words G is nilpotent of class 2. If λ, μ are any integers then

$$[a_i^\lambda, b_i^\mu] = c_i^{\lambda\mu}.$$

Hence a_i (and also b_i) has order $p^{\varrho(i)}$ modulo the centre. With the notation

$$F_i = N_i Z / Z$$

of § 5, each F_i is a direct product

$$F_i = \{Z a_i\} \times \{Z b_i\}$$

of two cycles of order $p^{\varrho(i)}$; hence G/Z is the restricted direct product of cyclic groups.

We denote by A the subgroup of G generated by all a_i and by K the group generated by A and C:

$$A = \{a_1, a_2, \dots\}, \qquad K = \{A, C\} = AC.$$

These groups are abelian.

Lemma 6.2. *If*

$$a_1^{\lambda(1)} a_2^{\lambda(2)} \dots a_q^{\lambda(q)} = c \in C,$$

then $\lambda(i) \equiv 0 \bmod p^{\varrho(i)}$ $(i = 1, \dots, q)$.

Proof. Transforming by b_i we get

$$b_i^{-1} c\, b_i = c \cdot c_i^{\lambda(i)}.$$

But $c \in C \subseteq Z$; hence $c^{\lambda(i)} = 1$, and $\lambda(i)$ is divisible by the order $p^{\varrho(i)}$ of c_i.

Lemma 6.3. *Put*

$$L = \{a_1, \ldots, a_s, a_{s+1}^p, a_{s+2}^p, \ldots, a_{s+t}^p, a_{s+t+1}, a_{s+t+2}, \ldots, c_1, c_2, \ldots\};$$

let R be a subgroup of L and $Q = \{R, a_{s+1}, a_{s+2}, \ldots, a_{s+t}\}$. Then Q/R can not be generated by fewer than t elements.

Proof. It clearly suffices to prove this in the special case that $R = L$, when $Q = K$; for in any case K/L is isomorphic to a factor group of Q/R. Now K/L is an elementary abelian group of order p^{t-s}; for if

$$(L\, a_{s+1})^{\lambda(1)} (L\, a_{s+2})^{\lambda(2)} \ldots (L\, a_{s+t})^{\lambda(t)} = L$$

is a relation between the $t-s$ obvious generators of K/L, then by an easy application of Lemma 6.2

$$\lambda(1) \equiv \lambda(2) \equiv \cdots \equiv \lambda(t) \equiv 0 \pmod{p},$$

that is to say, the relation is trivial; and the lemma follows.

Lemma 6.4. *Let R be a subgroup of $\{a_1, a_2, \ldots, a_s\}$, and let $k_1, k_2, \ldots, k_u$ be u elements of order p in K but not in R. Let t be such that*

$$u < \frac{p^t - 1}{p - 1}$$

(thus $t = [\log_2 u] + 2$ will do for all p). Then there are numbers $\lambda(1), \lambda(2), \ldots, \lambda(t)$, not all divisible by p, such that the group $\{R, a^*\}$ obtained from R by adjoining the element

$$a^* = a_{s+1}^{\lambda(1)} a_{s+2}^{\lambda(2)} \ldots a_{s+t}^{\lambda(t)}$$

also contains none of $k_1, k_2, \ldots, k_u$.

Proof. We put $Q = \{R, a_{s+1}, a_{s+2}, \ldots, a_{s+t}\}$, and consider maximal cyclic subgroups of Q/R. If $\{R\, a^*\}$ is such a maximal cyclic subgroup, and if

$$a^* = a_{s+1}^{\lambda(1)} a_{s+2}^{\lambda(2)} \ldots a_{s+t}^{\lambda(t)},$$

then $\lambda(1), \lambda(2), \ldots, \lambda(t)$ can not all be divisible by p; for if $\lambda(i) = p\,\mu(i)$, and if

$$a^{**} = a_{s+1}^{\mu(1)} a_{s+2}^{\mu(2)} \ldots a_{s+t}^{\mu(1)},$$

then $\{Ra^{**}\}$ contains $\{Ra^*\}$ properly; this follows from the fact that the order of a^{**} modulo the centre Z of G is a power of p. If there is a maximal cyclic subgroup $\{Ra^*\}$ of Q/R which contains none of $Rk_1, Rk_2, \ldots, Rk_u$, then $\{R, a^*\}$ does not contain any of $k_1, k_2, \ldots, k_u$, and the lemma follows. It remains to show that such a maximal cyclic subgroup exists. If Q/R is infinite, then—being finitely generated—it contains infinite maximal cyclic subgroups, and any one of these will do, as it can not contain any element of order p. Assume then that Q/R is finite. It can evidently be generated by t elements and—by Lemma 6.3—by no fewer than t elements. Hence it is the direct product of t cyclic subgroups; these all have order divisible by p.

The elements of order p in Q/R therefore form a subgroup of order p^t; this consists of

$$\frac{p^t-1}{p-1} > u$$

cycles of order p. Thus there is at least one such cycle containing none of $Rk_1, Rk_2, \ldots, Rk_u$; and a maximal cyclic subgroup of Q/R containing such a cyclic subgroup of order p then has the required property. This completes the proof of the lemma.

Corollary 6.5. *In the lemma, the assumption on the order of* $k_1, k_2, \ldots, k_u$ *can be relaxed to require them only to be elements of finite orders all divisible by* p.

For then certain powers of them have order p, and a subgroup not containing these powers can not contain the elements themselves, either.

§ 7. The class [*II*].

We denote by $[II]$ the class of those H-groups G that satisfy, in addition to (5.1), the conditions

$$c_1 = c_2 = \ldots = c, \text{ say}, \qquad c^p = 1. \tag{7.1}$$

In such a group the commutator c lies in the intersection of several (in fact all) N_i and, therefore, in the centre; it follows that the conditions (6.1) are also satisfied, with $\varrho(1) = \varrho(2) = \cdots = 1$, and

$$[II] \subseteqq [I].$$

The derived group of a II-group is the cycle $\{c\}$ of order p. Hence the group is an FD-group, and using (5.2) we see that

$$[II] \subseteqq [FD] - [FIZ]. \tag{7.2}$$

The existence of II-groups is easily verified; examples are the restricted direct product of infinitely many quaternion groups with amalgamated centre [*10*, p. 238] or the similarly constructed groups of Erdös [*4*, p. 58].

We further denote by $[II^*]$ the class of groups

$$G^* = \{c^*, a_1^*, b_1^*, a_2^*, b_2^*, \ldots\}$$

subject to

$$\begin{cases} [a_i^*, b_i^*] = c^* \neq 1, \\ [a_i^*, a_j^*] = [a_i^*, b_j^*] = [b_i^*, b_j^*] = 1, \\ c^{*p} = 1, \end{cases} \tag{7.3}$$

(where again $i, j = 1, 2, \ldots$ and $i \neq j$) and, moreover,

$$c^* \notin \{a_1^*, a_2^*, \ldots\}.$$

The first three conditions ensure that such a group G^* is a II-group.

Lemma 7.4. *Every group* $G \in [II]$ *contains a subgroup* $G^* \in [II^*]$.

Proof. We construct G^* by an inductive procedure. Put $c^* = c$, $G_0^* = A_0^* = \{c^*\}$, and assume we have already defined

$$G_n^* = \{c^*, a_1^*, b_1^*, a_2^*, b_2^*, \ldots, a_n^*, b_n^*\},$$
$$A_n^* = \{a_1^*, a_2^*, \ldots, a_n^*\}$$

such that (7.3) is satisfied for $i, j = 1, \ldots, n$, $i \neq j$, that $c^* \notin A_n^*$, and that

$$A_n^* \leqq A_{2n} = \{a_1, a_2, \ldots, a_{2n}\}.$$

We apply Lemma 6.4, with $R = A_n^*$, $s = 2n$, $u = 1$, $t = 2$, and $k_1 = c^*$, and find an element

$$a_{n+1}^* = a_{2n+1}^{\lambda(1)} a_{2n+2}^{\lambda(2)}$$

such that

$$c^* \notin A_{n+1}^* = \{a_1^*, a_2^*, \ldots, a_{n+1}^*\}.$$

Not both of $\lambda(1), \lambda(2)$ are divisible by p. Let $\lambda(z)$, where $z = 1$ or 2, be prime to p, and let μ be its reciprocal modulo p:

$$\lambda(z) \cdot \mu \equiv 1 \pmod{p}.$$

Put

$$b_{n+1}^* = b_{2n+z}^{\mu}.$$

Then $[a_{n+1}^*, b_{n+1}^*] = c^*$. It follows that $G_{n+1}^* = \{c^*, a_1^*, b_1^*, \ldots, a_{n+1}^*, b_{n+1}^*\}$ and A_{n+1}^* have again all the required properties. Finally we put

$$G^* = \bigcup_{n=1}^{\infty} G_n^*;$$

then $G^* \in [II^*]$, and the lemma follows.

Lemma 7.5. *If* $G^* \in [II^*]$ then $G^* \notin [X]$.

Proof. We show that all the conjugates $b_i^{*-1} A^* b_i^*$ are distinct. Let $i \neq j$; then $b_i^{*-1} A^* b_i^*$ contains

$$b_i^{*-1} a_i^* b_i^* = a_i^* c^*$$

whereas $b_j^{*-1} A^* b_j^*$ contains

$$b_j^{*-1} a_i^* b_j^* = a_i^*.$$

If $b_i^{*-1} A^* b_i^*$ coincided with $b_j^{*-1} A^* b_j$, then $c^* \in b_i^{*-1} A^* b_i^*$ and also

$$c^* = b_i^* c^* b_i^{*-1} \in b_i^* (b_i^{*-1} A^* b_i^*) b_i^{*-1} = A^*.$$

But this contradicts the definition of a II^*-group. Thus A^* has infinitely many distinct conjugates, and the lemma follows.

Corollary 7.6. $[II] \cap [X] = \emptyset$.

This follows from Lemmas 7.4, 7.5, and 2.2.

§ 8. The class [*III*].

We denote by $[III]$ the class of those I-groups G that satisfy, in addition to (5.1) and (6.1), the conditions

(8.1) $$c_i \neq c_j \quad \text{when} \quad i \neq j.$$

In such a group the derived group

$$C = \{c_1, c_2, \ldots\}$$

is infinite, and using (5.2) we see that

$$[III] \subseteq [FC] - [FD].$$

The existence of such groups is easily verified: the restricted direct product of infinitely many quaternion groups may serve as an example.

Lemma 8.2. *If $G \in [III]$ then G is not a BFC-group.*

Proof. The element $a = a_1 a_2 \ldots a_q$ has at least $q+1$ conjugates, namely apart from itself the q elements

$$b_i^{-1} a\, b_i = a\, c_i \qquad (i = 1, 2, \ldots, q).$$

These are all distinct. As q can be taken arbitrarily large, the cardinals of classes of conjugate elements are unbounded, and the lemma follows.

Next we denote by $[III^*]$ the class of groups $G \in [III]$ that satisfy, moreover,

$$c_i^{-1} c_j \notin A = \{a_1, a_2, \ldots\} \tag{8.3}$$

for $i, j = 1, 2, \ldots$ and $i \neq j$.

In such a group at most one c_i can belong to A, and we may assume that it is c_1, if any: then $c_i \notin A$ for $i = 2, 3, \ldots$.

Lemma 8.4. *A III^*-group G is neither an X-group nor a Y-group.*

Proof. We first show that all the conjugates $b_i^{-1} A\, b_i$ are distinct. Assume on the contrary that $i \neq j$ but $b_i^{-1} A\, b_i = b_j^{-1} A\, b_j$. Then both the elements

$$b_i^{-1} a_i b_i = a_i c_i,$$
$$b_j^{-1} a_i b_j = a_i$$

are contained in $b_i^{-1} A\, b_i$, and so then is c_i. Similarly $c_j \in b_i^{-1} A\, b_i$, and then

$$c_i^{-1} c_j = b_i c_i^{-1} c_j b_i^{-1} \in b_i (b_i^{-1} A\, b_i)\, b_i^{-1} = A,$$

contrary to (8.3). Thus A has infinitely many distinct conjugates, and $G \notin [X]$. Next, a normal subgroup H of G that contains A must also contain all commutators $[a_i, b_i] = c_i$, but by (8.3) no two of these are congruent modulo A, and $|\, H : A \,|$ is infinite. It follows that $G \notin [Y]$. This completes the proof of the lemma.

Lemma 8.5. *Every group $G \in [III]$ contains a subgroup $G^* \in [III^*]$.*

Proof. The plan of the proof is as for Lemma 7.4, but it requires more elaborate arguments. The generators of the group G^* to be constructed will be denoted by $a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots$, and conditions like (5.1), (6.1), when applied to these generators, will also be given an asterisk. Assume the groups

$$G_n^* = \{a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots, a_n^*, b_n^*, c_n^*\},$$
$$A_n^* = \{a_1^*, a_2^*, \ldots, a_n^*\},$$
$$C_n^* = \{c_1^*, c_2^*, \ldots, c_n^*\}$$

are defined and satisfy (5.1*), (6.1*), (8.1*), and (8.3*) as far as applicable; and assume further that there is a number $s = s(n)$ such that

$$A_n^* \subseteq A_s = \{a_1, a_2, \ldots, a_s\},$$
$$C_n^* \subseteq C_s = \{c_1, c_2, \ldots, c_s\},$$

and if $i \neq j$ and $c_i^{-1} c_j \in A_n^*$ then $i \leq s$ and $j \leq s$. We then determine t such that

$$\frac{p^t - 1}{p - 1} > (n + t)(n + t - 1)$$

and put $(n+t)(n+t-1) = u$. We further put

$$\varsigma_1 = c_1^*, \varsigma_2 = c_2^*, \ldots, \varsigma_n = c_n^*, \varsigma_{n+1} = c_{s+1}, \varsigma_{n+2} = c_{s+2}, \ldots, \varsigma_{n+t} = c_{s+t}$$

and denote the $(n+t)(n+t-1) = u$ quotients $\varsigma_i^{-1} \varsigma_j$ with $i \neq j$ by $k_1, k_2, \ldots, k_u$. None of these is contained in A_n^*. We can, therefore, apply Lemma 6.4 and Corollary 6.5, with $R = A_n^*$, and find an element

$$a^* = a_{s+1}^{\lambda(1)} a_{s+2}^{\lambda(2)} \ldots a_{s+t}^{\lambda(t)}$$

such that $\{A_n^*, a^*\}$ contains none of the elements $k_1, k_2, \ldots, k_u$. Moreover at least one of the exponents $\lambda(1), \ldots, \lambda(t)$, let us say $\lambda(z)$, is not divisible by p. Let μ be its reciprocal modulo the order $p^{\varrho(s+z)}$ of c_{s+z}, that is

$$\lambda(z) \cdot \mu \equiv 1 \pmod{p^{\varrho(s+z)}}.$$

We put

$$a_{n+1}^* = a^*,$$
$$b_{n+1}^* = b_{s+z}^{\mu}$$

and

$$[a_{n+1}^*, b_{n+1}^*] = c_{s+z} = c_{n+1}^*.$$

Then

$$G_{n+1}^* = \{G_n^*, a_{n+1}^*, b_{n+1}^*, c_{n+1}^*\},$$
$$A_{n+1}^* = \{A_n^*, a_{n+1}^*\},$$
$$C_{n+1}^* = \{C_n^*, c_{n+1}^*\}$$

satisfy (5.1*), (6.1*), (8.1*), (8.3*) as far as applicable. Moreover, as A_{n+1}^* is a finitely generated abelian group, it can contain only a finite number of quotients $c_i^{-1} c_j$ for $i \neq j$. There is thus an integer $s' = s(n+1)$ such that if $i \neq j$ and $c_i^{-1} c_j \in A_{n+1}^*$, then $i \leq s'$, $j \leq s'$, and such that at the same time $A_{n+1}^* \subseteq A_{s'}$ and $C_{n+1}^* \subseteq C_{s'}$. Thus the procedure works inductively; to start it we put $G_0^* = A_0^* = C_0^* = \{1\}$ and $s(0) = 0$. Finally we put

$$G^* = \bigcup_{n=1}^{\infty} G_n^*;$$

then $G^* \in [III^*]$ and the lemma follows.

Corollary 8.6. $[III] \cap ([BFC] \cup [X] \cup [Y]) = \emptyset$.

This follows from Lemmas 8.2, 8.4, 8.5, and 2.2.

§ 9. The class [J].

We denote by $[J]$ the class of H-groups G that satisfy, in addition to (5.1), the relations

$$\begin{cases} [a_i, c_i] = [b_i, c_i] = 1, \\ c_i^{p(i)} = 1, \end{cases} \tag{9.1}$$

where $i = 1, 2, \ldots$ and $p(1), p(2), \ldots$ are different prime numbers.

Again there may be further relations, but they must be compatible with $c_i \neq 1$ for all i. In a J-group G the subgroup

$$C = \{c_1, c_2, \ldots\}$$

is the derived group; it is contained in the centre, in other words G is nilpotent of class 2. If λ, μ are any integers, then

$$[a_i^\lambda, b_i^\mu] = c_i^{\lambda\mu}.$$

Hence a_i (and also b_i) has order $p(i)$ modulo the centre.

Lemma 9.2. *If $G \in [J]$ then G is not a BFC-group.*

Proof. The element $a = a_1 a_2 \ldots a_q$ has at least $q+1$ conjugates, namely apart from itself the q elements

$$b_i^{-1} a b_i = a c_i \qquad (i = 1, 2, \ldots, q).$$

They are all distinct, because the c_i have different orders and are, therefore, distinct. As q can be taken arbitrarily large, the cardinals of classes of conjugate elements are unbounded, and the lemma follows.

Next we denote by $[J^*]$ the class of groups $G \in [J]$ that satisfy, moreover,

$$c_i^{-1} c_j \notin A = \{a_1, a_2, \ldots\} \tag{9.3}$$

for $i, j = 1, 2, \ldots$ and $i \neq j$.

Lemma 9.4. *A J^*-group G is neither an X-group nor a Y-group.*

The proof is the same as that of Lemma 8.4, with (9.3) substituted for (8.3).

Lemma 9.5. *Let N be a group generated by two elements a, b such that*

$$[a, b] = c \neq 1,$$
$$[a, c] = [b, c] = 1,$$
$$c^p = 1$$

where p is an odd prime. Then N has a subgroup $N^* = \{a^*, b^*\}$ such that the order of a^* is infinite or a power of p, and

$$[a^*, b^*] = c \notin \{a^*\}.$$

Proof. If a has infinite order we put $a^* = a$, $b^* = b$; if a has finite order and b infinite order, we put $a^* = b$, $b^* = a^{-1}$. There only remains the case that a, b have finite orders α, β respectively. The order of N then is finite, namely a divisor of $\alpha\beta p$. We choose numbers λ, μ such that the element

$$a^* = a^\lambda b^\mu$$

(i) has order a power of p, (ii) does not generate c (that is, $c \notin \{a^*\}$), (iii) has maximal order subject to (i), (ii). Then λ, μ are not both divisible by p; for assume on the contrary that

$$\lambda = p\,\lambda', \quad \mu = p\,\mu'$$

and put

$$a' = a^{\lambda'} b^{\mu'}.$$

Then—as p is an odd prime and therefore divides $\frac{1}{2}p(p-1)$—

$$a'^p = a^{p\lambda'} b^{p\mu'} c^{\frac{1}{2}p(p-1)\lambda'\mu'} = a^\lambda b^\mu = a^*,$$

and a' satisfies (i) and (ii) and has greater order than a^*, contrary to the choice of a^*. If λ is prime to p, we put $b^* = b^\nu$ where ν is the reciprocal of λ modulo p, that is $\lambda\nu \equiv 1 \pmod{p}$. If λ is divisible by p, we put $b^* = a^\pi$ where similarly $\mu\pi \equiv -1 \pmod{p}$. It is then easy to verify that $N^* = \{a^*, b^*\}$ has the required properties, and the lemma follows.

Lemma 9.6. *Every group $G \in [J]$ has a homomorphic image G^0 with a subgroup $G^* \in [J^*]$.*

Proof. We apply the isoclinic homomorphism of G onto G^0 described in the proof of Lemma 4.7; then G^0 is periodic by Corollary 4.8, and one easily verifies that G^0 is also a J-group. Next, if one of the primes $p(i)$ is 2, we omit the corresponding a_i, b_i, c_i from the generators. We then rename and renumber the generators so that

$$G^0 = \{a_1, b_1, c_1, a_2, b_2, c_2, \ldots\}.$$

In each of the groups $N_i = \{a_i, b_i\}$ we choose a subgroup $N_i^* = \{a_i^*, b_i^*\}$ according to Lemma 9.5; then a_i^*, being periodic, has order a power of $p(i)$, and $[a_i^*, b_i^*] = c_i \notin \{a_i^*\}$. We put

$$G^* = \{a_1^*, b_1^*, c_1, a_1^*, b_1^*, c_2, \ldots\}$$

and

$$A^* = \{a_1^*, a_2^*, \ldots\}, \qquad C = \{c_1, c_2, \ldots\}.$$

Then $K^* = CA^*$ is a periodic abelian group. Its primary components are either trivial, namely if the prime p does not occur among the $p(i)$, or the direct product

$$\{c_i\} \times \{a_i^*\},$$

if $p = p(i)$. As a periodic abelian group is the restricted direct product of its primary components, K^* is seen to be the restricted direct product of all cycles $\{c_i\}$ and $\{a_i^*\}$, and then also of C and A^*:

$$K^* = C \times A^*.$$

The defining property (9.3) of a J^*-group is a trivial consequence; thus $G^* \in [J^*]$, and the lemma follows.

Corollary 9.7. $\qquad [J] \cap ([BFC] \cup [X] \cup [Y]) = \emptyset.$

This follows from Lemmas 9.2, 9.4, 9.6, and 2.2.

§ 10. The class $[K]$.

We denote by $[K]$ the class of H-groups G in which distinct subgroups $N_i = \{a_i, b_i\}$ have trivial intersection:

$$N_i \cap N_j = \{1\} \quad \text{when} \quad i \neq j. \tag{10.1}$$

It follows immediately from the remarks at the beginning of § 5 that G then is the restricted direct product of $N_1, N_2, \ldots$. Conversely, if an infinite sequence of strictly non-abelian two-generator FC-groups is given, then their restricted direct product is easily seen to be a K-group.

The derived group G' of a K-group G is the restricted direct product of the derived groups $N_1', N_2', \ldots$, none of which is trivial: hence G' is infinite, and using (5.2) we see that

$$[K] \subseteqq [FC] - [FD].$$

It may be remarked that if every N_i, and therefore also G, is nilpotent of class 2, then G is also a III-group; but not every III-group is a K-group.

Lemma 10.2. *If $G \in [K]$ then G is not a BFC-group.*

This is proved as Lemma 8.2.

Next we denote by $[K^*]$ the class of groups $G \in [K]$ that satisfy, moreover,

$$c_i \notin \{a_i\} \tag{10.3}$$

for $i = 1, 2, \ldots$.

Such a group satisfies the apparently stronger condition

$$c_i^{-1} c_j \notin A = \{a_1, a_2, \ldots\} \tag{10.4}$$

for $i, j = 1, 2, \ldots$ and $i \neq j$. For every element of a K-group G is of the form

$$g = n_1 n_2 \ldots n_q, \qquad n_i \in N_i,$$

and this representation is unique apart from trivial factors that can be added or omitted with a change of q. Thus if $c_i^{-1} c_j \in A$, then $c_i^{-1} c_j = a_i^{\lambda} a_j^{\mu}$ and $c_i = a_i^{-\lambda}$, $c_j = a_j^{\mu}$, contrary to (10.3).

Lemma 10.5. *A K^*-group G is neither an X-group nor a Y-group.*

This follows from (10.4) in precisely the same way as Lemma 8.4 follows from (8.3).

Lemma 10.6. *Every group $G \in [K]$ contains a subgroup $G^* \in [K^*]$.*

Proof. If (for some value of i) $c_i \in \{a_i\}$, then b_i transforms a_i into a power of a_i. We therefore have to look for elements that are transformed into other elements not powers of themselves. A non-abelian group in which all the conjugates of every element are amongst its powers only, is hamiltonian; the direct product of two non-abelian groups is not hamiltonian. Hence in every product $N_{2i-1} \times N_{2i}$ we can select two elements a_i^*, b_i^*, say, such that b_i^* transforms a_i^* into an element that is not a power of a_i^*. Then

$$[a_i^*, b_i^*] = c_i^* \notin \{a_i^*\}.$$

Finally we put $G^* = \{a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots\}$. This is then a K^*-group, and the lemma follows.

Corollary 10.7. $[K] \cap ([BFC] \cup [X] \cup [Y]) = \emptyset$.

This follows from Lemmas 10.2, 10.5, 10.6, and 2.2.

§ 11. The class [*FC*] — [*FD*].

Throughout this section we assume that G is an FC-group but not an FD-group. By selecting suitable subgroups we show that G is then neither a BFC-group, nor an X-group, nor a Y-group.

Lemma 11.1. *Every group* $G \in [FC] - [FD]$ *contains a subgroup* $G^1 \in [H] - [FD]$.

Proof. We construct G^1 by an inductive process. Assume elements $a_1, b_1, c_1, \ldots, a_{n-1}, b_{n-1}, c_{n-1}$ so chosen in G that

$$\text{(11.2)} \qquad \begin{cases} [a_i, b_i] = c_i \neq c_j \neq 1, \\ [a_i, a_j] = [a_i, b_j] = [b_i, b_j] = 1 \end{cases}$$

for all i and $j \neq i$. This is vacuously true for $n = 1$. The centralizer $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ has finite index in G, and cannot be an FD-group, by Corollary 4.2. The derived group of $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ is then infinite, and we can find two elements a_n, b_n, say, in $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ whose commutator $c_n = [a_n, b_n]$ is different from $1, c_1, c_2, \ldots, c_{n-1}$. Then $a_1, b_1, c_1, \ldots, a_n, b_n, c_n$ clearly satisfy (11.2). Finally we put

$$G^1 = \{a_1, b_1, c_1, a_2, b_2, c_2, \ldots\};$$

then G^1 is seen to be an H-group, and the construction ensures that the commutators c_i are all different; hence $G^1 \notin [FD]$, and the lemma follows.

Lemma 11.3. *If either* $[a_i, c_i] \neq 1$ *or* $[b_i, c_i] \neq 1$ *for infinitely many values of* i, *then* G^1 *is neither a BFC-group, nor an X-group, nor a Y-group.*

Proof. Let G^2 denote the subgroup of G^1 that is generated by all $N_i = \{a_i, b_i\}$ $(= \{a_i, b_i, c_i\})$ in which a_i, b_i do not both commute with c_i. Then G^2 is again an H-group but not an FD-group. Denote the centre of G^2 by Z^2. Its factor group G^2/Z^2 is the restricted direct product of infinitely many finite non-abelian groups

$$F_i = N_i Z^2/Z^2,$$

and is, therefore, a K-group. Then by Corollary 10.7 and repeated application of Lemma 2.2, $G^1 \notin [BFC] \cup [X] \cup [Y]$, and the lemma follows.

Lemma 11.4. *If* $[a_i, c_i] = [b_i, c_i] = 1$ *for all but a finite number of values of* i, *then* G^1 *contains a subgroup* $G^3 \in [H] - [FD]$ *that is nilpotent of class* 2 *(or, in other words, all of whose commutators lie in the centre).*

This is trivial: we only have to omit from the set of generators those finitely many a_i, b_i, c_i for which $[a_i, c_i] \neq 1$ or $[b_i, c_i] \neq 1$.

We note that in such a nilpotent group of class 2

(11.5) $$[a, b] = c \quad \text{implies} \quad [a^\lambda, b^\mu] = c^{\lambda\mu}$$

—a fact that we have repeatedly used. Next we note that the derived group

$$C^3 = \{c_1, c_2, \ldots\}$$

of G^3 is, as a periodic abelian group, the restricted direct product of its primary components.

Lemma 11.6. *Let $G^3 \in [H] - [FD]$ be nilpotent of class 2. If its derived group C^3 has an infinite primary component, then G^3 contains a subgroup $G^4 \in [III]$.*

Proof. Let the primary p-component C_p^3 of C^3 be infinite. If the order $\gamma(i)$ of c_i is factorized in the form

$$\gamma(i) = p^{\varrho(i)}\gamma'(i)$$

with $\gamma'(i)$ prime to p, then

$$C_p^3 = \{c_1^{\gamma'(1)}, c_2^{\gamma'(2)}, \ldots\}.$$

Also, by (11.5)

$$c_i^{\gamma'(i)} = [a_i, b_i^{\gamma'(i)}].$$

Among the $c_i^{\gamma'(i)}$ there must be infinitely many that are different from one another; we choose accordingly an infinite sequence $i(1), i(2), \ldots$ such that when $j \neq k$ then

$$c_j^* = c_{i(j)}^{\gamma'(i(j))} \neq c_k^* = c_{i(k)}^{\gamma'(i(k))},$$

and put

$$a_j^* = a_{i(j)}, \qquad b_j^* = b_{i(j)}^{\gamma'(i(j))}.$$

Finally we put $G^4 = \{a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots\}$. It is then immediately verified that G^4 is a III-group, and the lemma follows.

Lemma 11.7. *Let $G^3 \in [H] - [FD]$ be nilpotent of class 2. If its derived group C^3 has no infinite primary component, then G^3 contains a subgroup $G^5 \in [J]$.*

Proof. As C^3 is infinite, it must have infinitely many non-trivial primary components. It follows that there are infinitely many prime numbers each of which divides the order $\gamma(i)$ of at least one c_i; and infinitely many different c_i are required for this . Let then $i(1) < i(2) < \cdots$ be a sequence of indices and $p(1), p(2), \ldots$ a sequence of distinct primes such that $p(j)$ divides $\gamma(i(j))$. We put $\gamma(i(j)) = p(j) \cdot \gamma'(i(j))$ and

$$c_j^* = c_{i(j)}^{\gamma'(i(j))}.$$

Then c_j^* has order $p(j)$. Next we note that by (11.5)

$$c_j^* = [a_{i(j)}, b_{i(j)}^{\gamma'(i(j))}].$$

We therefore put

$$a_j^* = a_{i(j)}, \qquad b_j^* = b_{i(j)}^{\gamma'(i(j))},$$

and finally $G^5 = \{a_1^*, b_1^*, c_1^*, a_2^*, b_2^*, c_2^*, \ldots\}$. One confirms immediately that G^5 is a J-group, and the lemma follows.

Corollary 11.8. *If $G^3 \in [H] - [FD]$ is nilpotent of class 2, then it is neither a BFC-group, nor an X-group, nor a Y-group.*

This follows from Lemmas 11.6, 11.7, 2.2 and Corollaries 8.6, 9.7.

Corollary 11.9. *If G is an FC-group but not an FD-group, then G is neither a BFC-group, nor an X-group, nor a Y-group.*

This follows from Lemmas 11.1, 11.3, 11.4, 2.2, and Corollary 11.8.

§ 12. The class $[FD] - [FIZ]$.

We now assume that G is an FD-group but not a FIZ-group and show that it is then not an X-group either.

Lemma 12.1. *Every group $G \in [FD] - [FIZ]$ contains a subgroup $G^1 \in [H]$.*

Proof. We construct G^1 by an inductive process. Assume elements $a_1, b_1, c_1, \ldots, a_{n-1}, b_{n-1}, c_{n-1}$ so chosen that

$$\begin{cases} [a_i, b_i] = c_i \neq 1, \\ [a_i, a_j] = [a_i, b_j] = [b_i, b_j] = 1 \end{cases} \tag{12.2}$$

for all i and $j \neq i$. This is vacuously true for $n = 1$. The centralizer $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ has finite index in G, and cannot be an FIZ-group, by Corollary 4.4. Hence $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ is *a fortiori* non-abelian, and we can find two elements a_n, b_n, say, in $C(a_1, b_1, \ldots, a_{n-1}, b_{n-1})$ whose commutator $c_n = [a_n, b_n]$ is different from 1. Then $a_1, b_1, c_1, \ldots, a_n, b_n, c_n$ clearly satisfy (12.2). Finally we put

$$G^1 = \{a_1, b_1, c_1, a_2, b_2, c_2, \ldots\};$$

this is clearly an H-group, and the lemma follows. It may be noted that G^1, as a subgroup of an FD-group, is itself an FD-group.

Lemma 12.3. *Every group $G^1 \in [FD] \cap [H]$ contains a subgroup $G^2 \in [II]$.*

Proof. As the derived group of G^1 is finite, only a finite number of the c_i can be distinct, and at least one of them, c say, occurs infinitely often. Let $c = c_{i(1)} = c_{i(2)} = \cdots$. Then c lies in the centre of G^1, being generated by $a_{i(1)}$, $b_{i(1)}$ as well as by $a_{i(2)}$, $b_{i(2)}$ (cf. a remark on the immediate consequences of the definition of an H-group in § 5). As before we then have

$$[a_{i(j)}^{\lambda}, b_{i(j)}^{\mu}] = c^{\lambda\mu}.$$

Let p be a prime number dividing the order γ of c, and let $\gamma = p\,\gamma'$. Then $c^* = c^{\gamma'}$ has order p. We put

$$a_j^* = a_{i(j)}, \quad b_j^* = b_{i(j)}^{\gamma'},$$

and note that

$$\begin{aligned} [a_j^*, b_j^*] &= c^* \neq 1, \\ [a_j^*, c^*] = [b_j^*, c^*] &= 1, \\ c^{*p} &= 1. \end{aligned}$$

Hence the group $G^2 = \{c^*, a_1^*, b_1^*, a_2^*, b_2^*, \ldots\}$ is a II-group, and the lemma follows.

Corollary 12.4. *If G is an FD-group but not an FIZ-group, then G is not an X-group either.*

This follows from Lemmas 12.1, 12.3, 2.2 and Corollary 7.6.

§ 13. Conclusion.

We are now in a position to prove the main results announced in the introduction.

Theorem 13.1. *The classes $[BFC]$, $[FD]$, $[Y]$, and $[BY]$ all coincide.*

Proof. By Lemmas 2.3, 3.7

$$[FD] \subseteqq [BY] \subseteqq [Y] \subseteqq [FC].$$

Moreover $[FD] \subseteqq [BFC] \subseteqq [FC]$, as the order of the derived group is easily seen to provide a bound for the number of distinct conjugates of an element[5]). On the other hand, by Corollary 11.9,

$$[FC] - [FD] \subseteqq [FC] - ([BFC] \cup [Y]),$$

whence

$$[FD] = [BY] = [Y] = [BFC],$$

as asserted.

Theorem 13.2. *The classes $[FIZ]$, $[X]$, and $[BX]$ coincide.*

Proof. By Lemmas 2.3, 3.2

$$[FIZ] \subseteqq [BX] \subseteqq [X] \subseteqq [FC].$$

On the other hand, by Corollary 11.9

$$[FC] - [FD] \subseteqq [FC] - [X],$$

and by Corollary 12.4 also

$$[FD] - [FIZ] \subseteqq [FD] - [X].$$

Hence

$$[FIZ] = [BX] = [X],$$

as asserted.

It may be remarked that our methods do not allow us to state a bound for the order of the derived group in a group in which a finite bound is given for the cardinals of the classes of conjugate elements, or for the indices of subgroups in their normal closures; nor can we derive a bound for the index of the centre in a group in which a finite bound is given for the cardinals of the classes of conjugate subgroups. One is naturally led to conjecture that such bounds for the order of the derived group or the index of the centre do exist; and in the case of a given bound for the cardinals of the classes of conjugates elements this has very recently been confirmed by James Wiegold [*11*].

[5]) Cf. [*9*], proof of Theorem 5.4.

References.

[1] Baer, R.: Finiteness properties of groups. Duke Math. J. **15**, 1021—1032 (1948). — [2] Baer, R.: Das Hyperzentrum einer Gruppe. III. Math. Z. **59**, 299—338 (1953). — [3] Dedekind, R.: Über Gruppen, deren sämmtliche Teiler Normalteiler sind. Math. Ann. **48**, 548—561 (1897). — [4] Erdös, J.: The theory of groups with finite classes of conjugate elements. Acta Math. Acad. Sci. Hungar. **5**, 45—58 (1954). — [5] Haimo, F.: Groups with a certain condition on conjugates. Canad. J. Math. **4**, 369—372 (1952). — [6] Haimo, F.: The *FC*-chain of a group. Canad. J. Math. **5**, 498—511 (1953). — [7] Hall jr., M.: Subgroups of finite index in free groups. Canad. J. Math. **1**, 187—190 (1949). — [8] Hall, P.: The classification of prime-power groups. J. reine angew. Math. **182**, 130—141 (1940). — [9] Neumann, B. H.: Groups with finite classes of conjugate elements. Proc. London Math. Soc. (3) **1**, 178—187 (1951). — [10] Neumann, B. H.: Groups covered by permutable subsets. J. London Math. Soc. **29**, 236—248 (1954). — [11] Wiegold, J.: Unpublished.

Dept. of Math., The University, Manchester 13 (England).

(Eingegangen am 3. Januar 1955.)

Rogosinski, W. W.
Math. Zeitschr. Bd. 63, S. 97—108 (1955)

On finite systems of linear equations with an infinity of unknowns.

Dedicated to the memory of I. Schur, the great master and a helpful friend.

By
W. W. Rogosinski.

Introduction.

0.1. We are given n linear equations

$$\sum_{k=0}^{\infty} a_{ik} x_k = C_i \qquad (1 \leq i \leq n) \tag{0.1.1}$$

in an infinity of unknowns x_k. These and the given a_{ik} and C_i are complex (or real) numbers. The n sequences $\{a_{ik}\}$ are supposed to be linearly independent. It is further assumed that either, for a given p with $p \geq 1$,

$$\sum_{k=0}^{\infty} |a_{ik}|^p < \infty \qquad (1 \leq i \leq n); \tag{0.1.2}$$

or [this is the "convergency case" (c)] that

$$\lim_{k \to \infty} a_{ik} = a_{i0} \qquad (1 \leq i \leq n). \tag{0.1.3}$$

Note that, in the case (c), the first term a_{i0} of the sequence $\{a_{ik}\}$ is the limit of this sequence as $k \to \infty$.

0.2. As for the admissible solutions $\{x_k\}$, we demand, when $p > 1$, that

$$\sum_{k=0}^{\infty} |x_k|^{p'} < \infty \quad \text{where} \quad 1/p + 1/p' = 1; \tag{0.2.1}$$

when $p = 1$, that

$$\sup |x_k| < \infty; \tag{0.2.2}$$

and, in the case (c), that

$$\sum_{k=0}^{\infty} |x_k| < \infty. \tag{0.2.3}$$

In all cases, the left hand sides of (0.1.1) will then converge absolutely.

Amongst the infinity of solutions $\{x_k\}$ we consider *minimal solutions* $\{x_k^*\}$. These are solutions for which the left hand sides of (0.2.1), (0.2.2), or (0.2.3), respectively, become minimal. The existence of minimal solutions is in each case easy to prove. Also, if $p > 1$, the nature of the minimal solutions is known[1]). It can be expressed by the following theorem.

[1]) Riesz, F. [3], Chapter III, p. 42—77; in particular, p. 52—53, formula (21). The numbers refer to the list of references.

Theorem A. *If $p>1$, then the system* (0.1.1) *has exactly one minimal solution* $\{x_k^*\}$. *Moreover,*

$$x_k^* = \operatorname{sign} \overline{Q}_k \, |Q_k|^{p-1} \qquad (k \geq 0)^2), \tag{0.2.4}$$

where the Q_k are of the form

$$Q_k = \sum_{i=1}^{n} \Lambda_i a_{ik} \qquad (k \geq 0). \tag{0.2.5}$$

If $p=1$, and in the case (c), an infinity of minimal solutions will usually exist. The following results concerning their nature seem to be new.

Theorem B. *If $p=1$, then there exists at least one index K such that x_K^* is the same for every minimal solution* $\{x_k^*\}$ *of* (0.1.1). *Moreover,*

$$|x_K^*| = \max_k |x_k^*| \tag{0.2.6}$$

for every minimal solution.

In order to state the result in the case (c), we introduce the following terminology. A sequence $\{x_k\}$ is said to be *directed* if either (i) the x_k vanish eventually, or (ii) $\arg x_k \to \varphi \pmod{2\pi}$ as $k \to \infty$ for those k for which $x_k \neq 0$. Two sequences $\{x_k\}$ and $\{\xi_k\}$ are said to be *equally directed* if (α) each is directed, (β) $\arg x_k = \arg \xi_k$ whenever both $x_k \neq 0$ and $\xi_k \neq 0$, and (γ) both sequences have the same "direction" φ when they are in the above case (ii).

Theorem C. *In the case* (c) *every minimal solution* $\{x_k^*\}$ *of* (0.1.1) *is directed, and all are equally directed.*

In the case (c_0), when

$$\lim_{k \to \infty} a_{ik} = 0 \qquad (1 \leq i \leq n), \tag{0.2.7}$$

so that $a_{i0} = 0$, we can amplify Theorem C a little.

Theorem C_0. *In the case* (c_0) *all minimal solutions* $\{x_k^*\}$ *of* (0.1.1) *are equally directed. Moreover, there exists a K, the same for all minimal solutions, such that $x_k^* = 0$ when $k > K$.*

The 'real' case of (c), when all a_{ik}, C_i, and x_k are supposed to be real, deserves special mentioning. We shall say that two real sequences $\{x_k\}$ and $\{\xi_k\}$ have *equal sign* if, for every k, x_k and ξ_k are simultaneously non-negative, or non-positive. Theorem C then takes the following form.

Theorem C_r. *In the real case* (c) *all minimal solutions* $\{x_k^*\}$ *of* (0.1.1) *have equal sign. Moreover, there exists a K such that this sign remains constant when $k > K$.*

0.3. The preceding theorems may be illustrated by a simple geometrical situation in Euclidean space which they generalize. Consider, for instance, one real linear equation

$$a x + b y = C \tag{0.3.1}$$

---

[2]) We write $\operatorname{sign} 0 = 0$ and $\operatorname{sign} z = e^{i\varphi}$, when $\arg z = \varphi$.

in two real unknowns x and y. It may be interpreted as the equation of a line L in the Cartesian x, y-plane. We may also write (0.3.1) in 'infinite' form

$$a\,x + b\,y + o\cdot z + o\cdot s + \cdots = C. \tag{0.3.2}$$

If $p > 1$, it is clear, in view of (0.2.1), that we must have $z^* = s^* = \cdots = 0$ while $|x^*|^{p'} + |y^*|^{p'}$ must be made a minimum for every minimal solution. The simplest case $p = 2$, when $p' = 2$, is typical. The unique minimal solution is given by the point (x^*, y^*) on L nearest to the origin [compare Theorem A].

If $p = 1$, we must make $\max(|x^*|, |y^*|)$ a minimum μ, say. The $z^*, s^*, \ldots$ can be chosen arbitrarily but with absolute values at most μ. Geometrically, we have to find the square $|x| = \mu$, $|y| = \mu$ that touches the line L. There are two cases. Either L passes through exactly one vertex; then this vertex gives the unique minimal values x^* and y^*. Or L has one side in common with the square, when every point of this side gives a minimal pair (x^*, y^*). In either case, one of the coordinates at least is fixed for every minimal solution, with absolute value μ [compare Theorem B].

In the case (c_0), we must again have $z^* = s^* = \cdots = 0$ while $|x^*| + |y^*|$ becomes a minimum μ. This time we have to find the square $|x| + |y| = \mu$ that touches L. We have the same two cases as for $p = 1$. In the first, (x^*, y^*) is unique; in the second, all minimal pairs (x^*, y^*) belong to the same quadrant. In either case, Theorem C_0 is, clearly, satisfied.

0.4. The study of an infinite system of linearly independent linear equations

$$\sum_{k=0}^{\infty} a_{ik}\, x_k = C_i \qquad (i = 1, 2, \ldots) \tag{0.4.1}$$

can be attached, by well known arguments, to that of the finite case[3]. In the first place, of course, the system should have solutions. If so, then, for $p > 1$, there exists exactly one minimal solution, and an analogue of Theorem A remains valid. On the other hand, we shall see that neither Theorem B nor Theorem C holds in the infinite case except under further drastic assumptions.

0.5. The treatment of a finite system of linear equations in an infinity of unknowns in the older literature, as set out, for instance, in the book of F. Riesz [3], is direct and uses elementary algebraic and variational arguments. Prof. Rado has now also obtained a similar direct proof for our Theorem B. As for Theorem C, the property (β) above of all minimal solutions is quite easy to establish, and it should be possible to obtain also the properties (α) and (γ) by a direct method. My own proofs for all the above theorems are based on the methods of functional analysis applied to the appropriate sequential spaces to which the given $\{a_{ik}\}$ belong. In fact, these theorems turn out to be rather special applications of more general results concerning continuous linear functionals on these spaces, results which deserve some interest of their own. We require the well known form of these functionals and further the familiar Hahn-Banach extension theorem for functionals. It should be noted that the proof of this theorem requires no transfinite induction since

[3]) Riesz, F. [3], Chapter III.

all our spaces are separable. Our method of proof has the great advantage over the direct and apparently more elementary methods that it covers all cases uniformly in a short and straightforward way, and it seems to me the natural approach to these and many similar extremum problems. In any case, it is hoped that the results of this paper will add another example of the usefulness and power of abstract methods in quite concrete problems of analysis.

1. The general argument.

1.1. Let $\mathfrak{A}$ be a normed linear vector space over the field of complex (or real) numbers; $\mathfrak{A}$ need not be complete (a Banach space). If $a \in \mathfrak{A}$, the norm of a is denoted by $\|a\|$. The continuous (bounded) linear functionals (c.l.f.) on $\mathfrak{A}$ are denoted by $B(a)$; in the complex case they are complex-valued, in the real case real-valued. They form the conjugate space $\mathfrak{B}$ with norm $\|B\|$.

Suppose now that we are given n linearly independent points $a_i \in \mathfrak{A}$ and n arbitrary complex (real) numbers C_i. We may ask for all those B that satisfy

$$B(a_i) = C_i \qquad (1 \leq i \leq n). \tag{1.1.1}$$

Consider the n-dimensional subspace $\mathfrak{P}_n$ of $\mathfrak{A}$ of all linear combinations $p = p_n = \sum_1^n \lambda_i a_i$, where the λ_i are complex (real) numbers. Any solution of (1.1.1) must satisfy

$$B(p) = \sum_1^n \lambda_i B(a_i) = \sum_1^n \lambda_i C_i. \tag{1.1.2}$$

Conversely, the right hand side of (1.1.2) is a c.l.f. $\tilde{B}(p)$ defined on $\mathfrak{P}_n$ [4] which satisfies the equations

$$\tilde{B}(a_i) = C_i \qquad (1 \leq i \leq n). \tag{1.1.3}$$

The solutions B of (1.1.1) are, therefore, the 'extensions' of $\tilde{B}$ from $\mathfrak{P}_n$ to the whole of $\mathfrak{A}$, i.e. they are those B on $\mathfrak{A}$ for which $B(p) = \tilde{B}(p)$ for all $p \in \mathfrak{P}_n$. If $\|\tilde{B}\|_*$ denotes the norm of $\tilde{B}$ on $\mathfrak{P}_n$, then, clearly,

$$\|B\| \geq \|\tilde{B}\|_* \tag{1.1.4}$$

for any such extension. We now apply the HAHN-BANACH extension theorem [5] to the functionals $\tilde{B}$ on $\mathfrak{P}_n$. According to this theorem there exists at least one extension B^* of $\tilde{B}$ from $\mathfrak{P}_n$ to $\mathfrak{A}$ that is minimal in the sense that

$$\|B^*\| = \|\tilde{B}\|_*. \tag{1.1.5}$$

In view of (1.1.4) we call such a B^* a *minimal solution* for the system (1.1.1).

[4] See [6], p. 213.
[5] See [1], p. 55; and [2].

Next, it follows from the form (1.1.2) of $\widetilde{B}$, that there exists at least one point $P=\sum_1^n \Lambda_i a_i$ of $\mathfrak{P}_n$ such that $\|P\|=1$ and

$$\widetilde{B}(P)=\sum_1^n \Lambda_i C_i=\|\widetilde{B}\|_*. \tag{1.1.6}$$

We call P a *maximal point* for $\widetilde{B}$. By (1.1.5), we have

$$B^*(P)=\|B^*\|. \tag{1.1.7}$$

We note that this relation must hold for *every* pair of a minimal B^* and a maximal P. For any $p\in\mathfrak{P}_n$ with $\|p\|=1$ we have, of course, $|B^*(p)|\leq\|B^*\|$. *The discussion of equality here, as it holds in* (1.1.7), *allows in many cases additional information about the functional* B^*. This simple remark provides a powerful argument in applications.

We add the following remarks. The set of all minimal extensions B^* of $\widetilde{B}$ is a convex set in $\mathfrak{B}$. Similarly, the set of all maximal points P is convex in $\mathfrak{P}_n$, provided, however, that $\|\widetilde{B}\|_*>0$. In the "trivial" case when $\|\widetilde{B}\|_*=0$, all $p\in\mathfrak{P}_n$ with $\|p\|=1$ are maximal[6]).

A normed linear vector space $\mathfrak{V}$ is said to be *strictly convex* if, in the triangle relation

$$\|v_1+v_2\|\leq\|v_1\|+\|v_2\| \tag{1.1.8}$$

for the norm, equality occurs if, and only if, $\lambda_1 v_1=\lambda_2 v_2$ where $\lambda_1\geq 0$, $\lambda_2\geq 0$ and $\lambda_1+\lambda_2>0$.

If $\mathfrak{B}$ is strictly convex then the above convex set of the B^* reduces to one point: the minimal solution of (1.1.1) is then unique. Similarly, if $\mathfrak{P}_n$ is strictly convex, the maximal point P for $\widetilde{B}$ is unique, provided that $\|\widetilde{B}\|_*>0$[6]).

1.2. We consider, next, an infinite system of equations

$$B(a_i)=C_i \qquad (i\geq 1). \tag{1.2.1}$$

We assume that the sequence of points $a_i\in\mathfrak{A}$ is linearly independent. Let B_n^* be a minimal solution for the corresponding finite system (1.1.1). Clearly, the $\|B_n^*\|$ increase with n. It is well known that the system (1.2.1) has solutions $B\in\mathfrak{B}$ if, and only if,

$$\|B_n^*\|\uparrow M, \tag{1.2.2}$$

where M is finite[7]). Moreover, if this condition is satisfied, then there exists at least one minimal solution B^* of (1.2.1) for which $\|B^*\|=M$. For all other solutions, $\|B\|\geq M$. Again, the minimal solutions form a convex set in $\mathfrak{B}$, and if $\mathfrak{B}$ is strictly convex, then B^* is unique.

---

[6]) [5], p. 299—300.

[7]) [1], Theorem 5, p. 57 (for a real space).

Let $\mathfrak{P} = \bigcup_1^\infty \mathfrak{P}_n$ and let $\overline{\mathfrak{P}}$ be the closure of $\mathfrak{P}$ in $\mathfrak{A}$. If the unit sphere of $\overline{\mathfrak{P}}$ is weakly compact (or, equivalently, if $\overline{\mathfrak{P}}$ is reflexive), then there exists at least one maximal point $P \in \overline{\mathfrak{P}}$, such that $\|P\| = 1$ and

$$B^*(P) = \|B^*\|. \tag{1.2.3}$$

In this case we may again be able to draw some additional information concerning the minimal solutions B^*. It should, however, be noted, that without such or similar assumptions there will be, in general, no maximal points P to give us this information. Lastly, we remark that there exists at most one maximal point if $\overline{\mathfrak{P}}$ is strictly convex, provided that $\|B^*\| > 0$.

2. The space l^p, $p > 1$.

2.1. We consider the Banach space l^p, $p > 1$, of all complex (real) sequences $a = \{a_k\}$, $k \geqq 0$, with $\sum_{k=0}^\infty |a_k|^p < \infty$. The norm is $\|a\|_p = \left[\sum_{k=0}^\infty |a_k|^p\right]^{1/p}$.

It is well known that every c.l.f. b on l^p is of the form[8])

$$b(a) = \sum_{k=0}^\infty a_k x_k, \tag{2.1.1}$$

where

$$x = \{x_k\} \in l^{p'}, \quad 1/p + 1/p' = 1, \tag{2.1.2}$$

and $\|b\| = \|x\|_{p'}$. We call x the *kernel* of b; it is uniquely determined by b.

2.2. Let $\mathfrak{g}$ be a linear subspace of l^p. If $\tilde{b}$ is a c.l.f. defined on $\mathfrak{g}$, with norm $\|\tilde{b}\|_*$ on $\mathfrak{g}$, then, by the HAHN-BANACH extension theorem, there exists a representation

$$\tilde{b}(g) = \sum_{k=0}^\infty g_k x_k^* \qquad (g = \{g_k\} \in \mathfrak{g}) \tag{2.2.1}$$

where

$$x^* = \{x_k^*\} \in l^{p'} \quad \text{and} \quad \|x^*\|_{p'} = \|\tilde{b}\|_*. \tag{2.2.2}$$

In general, $\tilde{b}$ will also have other representations, with other kernels $x \in l^{p'}$. Then $\|x\|_{p'} \geqq \|\tilde{b}\|_*$ for all such kernels, so that x^* appears as a *minimal kernel* for $\tilde{b}$. Since the conjugate space $l^{p'}$ of l^p (the space of the c.l.f. on l^p) is strictly convex (by Minkowski's inequality), *the minimal kernel x^* is unique.*

2.3. A sequence $G = \{G_k\} \in \mathfrak{g}$ is said to be *maximal* for $\tilde{b}$ if

$$\text{(i)} \quad \|G\|_p = 1, \qquad \text{(ii)} \quad \tilde{b}(G) = \|\tilde{b}\|_*. \tag{2.3.1}$$

If $\|\tilde{b}\|_* = 0$, then all $G \in \mathfrak{g}$ with $\|G\|_p = 1$ are maximal. If $\|\tilde{b}\|_* > 0$ and $\mathfrak{g}$ is closed, then there exists exactly one maximal G. For, l^p is both reflexive

[8]) [1], p. 68 (for the real space l^p).

(its unit sphere is weakly compact) and strictly convex. In this case, we conclude from

$$\tilde{b}(G) = \sum_{k=0}^{\infty} G_k x_k^* = \|\tilde{b}\|_* = \|x^*\|_{p'} \tag{2.3.2}$$

that

$$x_k^* = \operatorname{sign} \overline{G}_k \, |G_k|^{p-1} \|\tilde{b}\|_* . \tag{2.3.3}$$

We thus obtain the following result.

Theorem 1. *Let $p > 1$, and let $\mathfrak{g}$ be a closed subspace of $\mathfrak{l}^p$. If $\tilde{b}$ is a continuous linear functional on $\mathfrak{g}$, then there exists a unique $\Gamma = \{\Gamma_k\} \in \mathfrak{g}$ such that*

$$\tilde{b}(g) = \sum_{k=0}^{\infty} g_k \operatorname{sign} \overline{\Gamma}_k \, |\Gamma_k|^{p-1}; \quad \|\tilde{b}\|_* = \|\Gamma\|_p^{p-1}. \tag{2.3.4}$$

Remarks. (i) If $\|\tilde{b}\|_* = 0$, then $\Gamma = 0$; otherwise $\Gamma = \|\tilde{b}\|_*^{1/p-1} G$. (ii) If $p = 2$, then there exists a unique $\Gamma \in \mathfrak{g}$ such that

$$\tilde{b}(g) = \sum_{k=0}^{\infty} g_k \overline{\Gamma}_k; \quad \|\tilde{b}\|_* = \|\Gamma\|_2. \tag{2.3.5}$$

(iii) If $\mathfrak{g} = \mathfrak{l}^p$, we find the existence of exactly one $A = A_k \in \mathfrak{l}^p$ such that

$$b(a) = \sum_{k=0}^{\infty} a_k \operatorname{sign} \overline{A}_k \, \|A_k\|^{p-1}; \quad \|b\| = \|A\|_p^{p-1}. \tag{2.3.6}$$

2.4. We use the notations in § 1. Let the $a_i = \{a_{ik}\}$, $1 \leq i \leq n$, $k \geq 0$, belong to $\mathfrak{A} = \mathfrak{l}^p$. By (2.1.1), the system of equations (1.1.1) in the unknown functionals $B = b$ takes the form of the system of linear equations (0.1.1) in the unknowns x_k. In fact, the conditions (0.1.2) and (0.2.1) state that $a_i \in \mathfrak{l}^p$ and $x = \{x_k\} \in \mathfrak{l}^{p'}$. We now apply Theorem 1 to the closed subspace $\mathfrak{g} = \mathfrak{P}_n$ spanned by the a_i and obtain Theorem A of the introduction.

We note also that Theorem A remains valid, with

$$x_k^* = \operatorname{sign} \overline{\Gamma}_k \, |\Gamma_k|^{p-1}, \quad \Gamma \in \overline{\mathfrak{P}}, \tag{2.4.1}$$

for an infinite system (0.4.1) of linear equations, provided that this system has at all solutions $x \in \mathfrak{l}^{p'}$. The necessary and sufficient condition for this is that the sequence of the $\|x^*(n)\|_p$ should remain bounded, where $x^*(n)$ is the minimal solution of the corresponding finite system (0.1.1).

3. The space $\mathfrak{l}$.

3.1. We consider the Banach space $\mathfrak{l}$ of all complex (real) sequenses $a = \{a_k\}$, $k \geq 0$, with $\sum_{k=0}^{\infty} |a_k| < \infty$. The norm is $\|a\| = \sum_{k=0}^{\infty} \|a_k\|$. Every c.l.f. b on $\mathfrak{l}$ can be represented in the form[9])

$$b(a) = \sum_{k=0}^{\infty} a_k x_k \tag{3.1.1}$$

[9]) [*1.*], p. 67 (for the real space $\mathfrak{l}$).

where

$$\|b\| = \sup |x_k| \quad (<\infty). \tag{3.1.2}$$

The *kernel* $x = \{x_k\}$ is uniquely determined by b.

3.2. Let $\mathfrak{g}$ be a linear subspace of $\mathfrak{l}$. If $\tilde{b}$ is a c.l.f. defined on $\mathfrak{g}$, with norm $\|\tilde{b}\|_*$ on $\mathfrak{g}$, then, by the Hahn-Banach theorem, there exists a representation

$$\tilde{b}(g) = \sum_{k=0}^{\infty} g_k x_k^* \qquad (g = \{g_k\} \in \mathfrak{g}) \tag{3.2.1}$$

where

$$\|\tilde{b}\|_* = \sup |x_k^*| . \tag{3.2.2}$$

Again $x^* = \{x_k^*\}$ appears as a *minimal kernel* for $\tilde{b}$ in the sense that, for every other representation of $\tilde{b}$ with kernel $x = \{x_k\}$, $\sup |x_k| \geq \|\tilde{b}\|_*$. The conjugate space of $\mathfrak{l}$ is the space $\mathfrak{l}^\infty$ of all bounded sequences $x = \{x_k\}$, with norm $\|x\| = \sup |x_k|$. It is not strictly convex. Hence, in general, we shall have an infinity of minimal kernels x^* for $\tilde{b}$. They form a convex set in $\mathfrak{l}^\infty$, lying on the sphere $\|x\| = \|\tilde{b}\|_*$.

3.3. A sequence $G = \{G_k\} \in \mathfrak{g}$ is said to be *maximal* for $\tilde{b}$ if

$$\text{(i)} \quad \|G\| = 1, \qquad \text{(ii)} \quad \tilde{b}(G) = \|\tilde{b}\|_*. \tag{3.3.1}$$

If $\|\tilde{b}\|_* = 0$, then all $G \in \mathfrak{g}$ with $\|G\| = 1$ are maximal. If $\|\tilde{b}\|_* > 0$, there may be no maximal sequences, even when $\mathfrak{g}$ is closed [since $\mathfrak{l}$ is not reflexive]; and there may be many [since $\mathfrak{l}$ is not strictly convex]. They form a convex set lying on the unit sphere $\|g\| = 1$ of $\mathfrak{g}$.

We now *assume* that $\tilde{b}$ has a maximal G. Then

$$\tilde{b}(G) = \sum_{k=0}^{\infty} G_k x_k^* = \|\tilde{b}\|_* = \sup |x_k^*| . \tag{3.3.2}$$

It, clearly, follows that

$$x_k^* = \operatorname{sign} \overline{G}_k \|\tilde{b}\|_* \quad \text{whenever} \quad G_k \neq 0. \tag{3.3.3}$$

We note that the set $K(G) = [k : G_k \neq 0]$ is not empty since $\|G\| = 1$. Also (3.3.3) must hold for every pair of a maximal G and a minimal x^*. It follows that, for all minimal kernels x^*, the x_k^* with k belonging to the non-empty "index set" $K = \bigcup_G K(G)$ are the same, and that for them $|x_k^*| = \|\tilde{b}\|_*$.

Theorem 2. *Suppose that $\mathfrak{g}$ is a subspace of $\mathfrak{l}$ and that the continuous linear functional $\tilde{b}$ on $\mathfrak{g}$ has a maximal sequence G. Then there exists a non-empty index set K and a uniquely determined sequence $\{\eta_k\}$ for $k \in K$, with $|\eta_k| = \|\tilde{b}\|_*$, such that every minimal representation of $\tilde{b}$ is of the form*

$$\tilde{b}(g) = \sum_{k \in K} g_k \eta_k + \sum_{k \notin K} g_k x_k^* \tag{3.3.4}$$

where $|x_k^*| \leq \|\tilde{b}\|_*$.

3.4. It follows from (3.3.4) that a necessary condition for $\tilde{b}$ to have a maximal G is that *all* minimal kernels x^* should satisfy

$$\sup|x_k^*| = \max|x_k^*| = \|\tilde{b}\|_*. \tag{3.4.1}$$

Conversely, suppose that *some* minimal kernel x^{**} satisfies (3.4.1). Let $K^* = [k:|x_k^{**}| = \|\tilde{b}\|_*]$. Then any sequence G with $\|G\| = 1$, satisfying

$$G_k = \begin{cases} 0 \\ 0 \end{cases} \quad \text{or} \quad x_k^{**} = \|\tilde{b}\|_* \operatorname{sign} \overline{G}_k \quad \text{when} \quad \begin{cases} k \in K^* \\ k \notin K^* \end{cases} \tag{3.4.2}$$

is, plainly, maximal, provided that $G \in \mathfrak{g}$; and these are all maximal G. Theorem 2 then holds and *all* minimal x^* satisfy (3.4.1). However, in general, such sequences G will not belong to $\mathfrak{g}$.

On the other hand, if $\mathfrak{g} = \mathfrak{l}$ [in which case the (minimal) kernel x of (3.1.1) is unique], then all the above G exist in $\mathfrak{l}$.

Theorem 3. *In order that the continuous linear functional b on $\mathfrak{l}$ should have a maximal sequence $A \in \mathfrak{l}$ it is necessary and sufficient that the kernel x of b should satisfy*

$$\sup|x_k| = \max|x_k| = \|b\|. \tag{3.4.3}$$

If $\tilde{K}$ is the non-empty index set $[k:|x_k| = \|b\|]$, then the maximal sequences A are exactly those, with $\|A\| = 1$, that satisfy

$$A_k = \begin{cases} 0 \\ 0 \end{cases} \quad \textit{or} \quad x_k = \|b\| \operatorname{sign} \overline{A}_k \quad \textit{when} \quad \begin{cases} k \in \tilde{K} \\ k \notin \tilde{K} \end{cases}. \tag{3.4.4}$$

3.5. If the $a_i = \{a_{ik}\}$, $1 \leq i \leq n$, $k \geq 0$, belong to $\mathfrak{A} = \mathfrak{l}$, that is, if they satisfy (0.1.2) for $p = 1$, then, by (3.1.1), the system (1.1.1) is equivalent to the system (0.1.1). Theorem 2, applied to the subspace $\mathfrak{P}_n$ of $\mathfrak{l}$ spanned by the a_i then yields Theorem B of the introduction.

The space $\mathfrak{l}$ is not reflexive. Hence for an infinite system (0.4.1) of linear equations, even if they have solutions, an "index theorem" will, in general, only hold if the corresponding c.l.f. $\tilde{b}$ has a maximal point in $\overline{\mathfrak{P}}$. This, for instance, will be the case when $\overline{\mathfrak{P}}$ itself is reflexive. In any case, the condition (3.4.1) for every minimal solution is a necessary condition for the existence of such a maximal point. If $\overline{\mathfrak{P}} = \mathfrak{l}$, it is also sufficient.

As an example, where no index theorem applies, consider the infinite system

$$\begin{cases} 0 \cdot x_0 + \frac{1}{2}\lambda_1(x_1 + x_2) + 0 \cdot (x_3 + x_4) + 0 \cdot (x_5 + x_6) + \cdots = 1 \\ 0 \cdot x_0 + 0 \cdot (x_1 + x_2) + \frac{1}{2}\lambda_2(x_3 + x_4) + 0 \cdot (x_5 + x_6) + \cdots = 1 \\ 0 \cdot x_0 + 0 \cdot (x_1 + x_2) + 0 \cdot (x_3 + x_4) + \frac{1}{2}\lambda_3(x_5 + x_6) + \cdots = 1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{cases} \tag{3.5.1}$$

where $\lambda_k > 1$, $k \geq 1$, and $\lambda_k \downarrow 1$. There exist solutions, for instance, $\tilde{x}_{2k} = \tilde{x}_{2k-1} = 1/\lambda_k$, $k \geq 1$, and $\tilde{x}_0$ arbitrary. Also, clearly, at least one of the two unknowns x_{2k-1}, x_{2k} must be not less than $1/\lambda_k$. Hence $\sup|x_k^*| = 1$ for any minimal solution (for instance, for the above $\tilde{x}$ if $|\tilde{x}_0| \leq 1$). This allows for some

variation in each couple x^*_{2k-1}, x^*_{2k} while x^*_0 is only restricted by $|x^*_0| \leq 1$. Hence there is no index K for which x^*_K must be the same for every minimal solution. Also there exist some minimal solutions for which $|x^*_k| < 1$ for all k.

4. The space c.

4.1. We consider the Banach space c of all complex (real) sequences $a = \{a_k\}$, $k \geq 1$, for which $a_k \to a_0$. The norm is $\|a\| = \sup |a_k|$. Every c.l.f. b on c can be represented in the form[10])

$$b(a) = \sum_{k=0}^{\infty} a_k x_k \tag{4.1.1}$$

where

$$\|b\| = \sum_{k=0}^{\infty} |x_k| = \|x\|_1. \tag{4.1.2}$$

The *kernel* $x = \{x_k\} \in \mathfrak{l}$ and is unique.

4.2. Let g be a linear subspace of c. If $\tilde{b}$ is a c.l.f. defined on g, with norm $\|\tilde{b}\|_*$ on g, then, by the Hahn-Banach theorem, there exists a representation

$$\tilde{b}(g) = \sum_{k=0}^{\infty} g_k x^*_k \quad (g = \{g_k\} \in \mathfrak{g}) \tag{4.2.1}$$

where

$$\|\tilde{b}\|_* = \sum_{k=0}^{\infty} |x^*_k| = \|x^*\|_1. \tag{4.2.2}$$

For any other representation of $\tilde{b}$ with kernel x we have $\|x\|_1 \geq \|\tilde{b}\|_*$ so that again x^* is a *minimal* kernel for $\tilde{b}$. The conjugate space $\mathfrak{l}$ of c is not strictly convex. Hence, in general, we shall have an infinity of minimal kernels x^* for $\tilde{b}$. They form a convex set lying on the sphere $\|x\|_1 = \|\tilde{b}\|_*$ in $\mathfrak{l}$.

4.3. A sequence $G = \{G_k\} \in \mathfrak{g}$ is *maximal* for $\tilde{b}$ if

$$\text{(i)} \quad \|G\| = 1, \qquad \text{(ii)} \quad \tilde{b}(G) = \|\tilde{b}\|_*. \tag{4.3.1}$$

If $\|\tilde{b}\|_* = 0$, then all $G \in \mathfrak{g}$ with $\|G\| = 1$ are maximal. If $\|\tilde{b}\|_* > 0$, there may be no maximal sequences, even when g is closed [since c is not reflexive]; and there may be many [since c is not strictly convex]. They form a convex set lying on the unit sphere $\|g\| = 1$ of g.

We now *assume* that $\tilde{b}$ has a maximal G. Then

$$\tilde{b}(G) = \sum_{k=0}^{\infty} G_k x^*_k = \|\tilde{b}\|_* = \sum_{k=0}^{\infty} |x^*_k|. \tag{4.3.2}$$

This implies

$$x^*_k = \begin{cases} 0 \\ 0 \end{cases} \quad \text{or} \quad \operatorname{sign} x^*_k = \operatorname{sign} \overline{G}_k \quad \text{whenever} \quad \begin{cases} |G_k| = 1 \\ |G_k| < 1 \end{cases}. \tag{4.3.3}$$

Hence either $x^*_k = 0$ eventually, or $\operatorname{sign} x^*_k \to \operatorname{sign} \overline{G}_0$ as $k \to \infty$ for the $x^*_k \neq 0$, where $G_0 = \lim G_k$: the sequence $x^* = \{x^*_k\}$ is directed. We also note that,

[10]) [1], p. 67 (for the real space c).

when $\|\tilde{b}\|_* > 0$, we must have at least one $x_k^* \neq 0$ and hence the corresponding $|G_k| = 1$. Also (4.3.3) must hold for every pair of a maximal G and a minimal x^*. Hence all minimal x^* are equally directed.

Theorem 4. *Suppose that* $\mathfrak{g}$ *is a subspace of* $\mathfrak{c}$ *and that the continuous linear functional* $\tilde{b}$ *on* $\mathfrak{g}$ *has a maximal sequence* G. *Then all minimal kernels* x^* *of* $\tilde{b}$ *are equally directed.*

Remarks. (i) If $\mathfrak{g} < \mathfrak{c}_0$ where $\mathfrak{c}_0$ is the subspace of all sequences g with $g_k \to 0$, then $G_0 = 0$. Hence there exists a K such that for all (equally directed) minimal kernels x^* of $\tilde{b}$

$$x_k^* = 0 \quad \text{for} \quad k > K. \tag{4.3.4}$$

(ii) In the case of the space $\mathfrak{c}_r$ of real convergent sequences, $|G_0| < 1$ or $G_0 = \pm 1$. It follows that there exists a K such that, for $k > K$, all minimal solutions x^* of $\tilde{b}$ have constant sign (non-negative, or non-positive).

4.4. Suppose that *some* minimal kernel x^{**} of $\tilde{b}$ is directed. Then any sequence G satisfying $\|G\| = 1$ and

$$|G_k| = 1, \quad \operatorname{sign} \overline{G}_k = \operatorname{sign} x_k^{**} \quad \text{whenever} \quad x_k^{**} \neq 0, \tag{4.4.1}$$

is, plainly, maximal for $\tilde{b}$ provided that $G \in \mathfrak{g}$. Theorem 4 then holds, and *all* x_k^* are equally directed and satisfy the analogue of (4.4.1). However, such $G \in \mathfrak{g}$ will not always exist. On the other hand, if $\mathfrak{g} = \mathfrak{c}$ (in which case the (minimal) kernel x of (4.1.1) is unique), then all these G are in $\mathfrak{c}$.

Theorem 5. *In order that the continuous linear functional* b *on* $\mathfrak{c}$ *should have a maximal sequence* $A \in \mathfrak{c}$ *it is necessary and sufficient that the kernel* x *of* b *should be directed.*

The maximal sequences A *are then exactly those that satisfy* $\|A\| = 1$ *and*

$$|A_k| = 1, \qquad \operatorname{sign} \overline{A}_k = \operatorname{sign} x_k \quad \textit{whenever} \quad x_k \neq 0. \tag{4.4.2}$$

Remarks. (i) For the space $\mathfrak{c}_0$ the characteristic condition is that the kernel x of b should be "finite"; i.e., that $x_k = 0$ eventually. This kernel (normalized by $x_0 = 0$) is unique.

(ii) For the real space $\mathfrak{c}_r$ the condition is that the kernel x of b should be of constant sign eventually.

4.5. If the $a_i = \{a_{ik}\}$ in (0.1.1) belong to $\mathfrak{c}$, then the application of Theorem 4 to the subspace $\mathfrak{P}_n$ of $\mathfrak{c}$ spanned by the a_i yields Theorem C of the introduction and its corollaries Theorems C_r and C_0.

The space $\mathfrak{c}$ is not reflexive. Hence for an infinite system (0.4.1) of linear equations, even if they have solutions, a "direction theorem' will usually only hold if the corresponding c.l.f. $\tilde{b}$ has a maximal sequence in $\overline{\mathfrak{P}}$. For the existence of a maximal sequence it is necessary that all minimal solutions should be equally directed. If $\overline{\mathfrak{P}} = \mathfrak{c}$, this is also sufficient.

The system of equations (in $\mathfrak{c}_0$)

$$\left\{\begin{array}{l} 1\cdot x_1+0\cdot x_2+0\cdot x_3+0\cdot x_4+\cdots=1 \\ 0\cdot x_1-4\cdot x_2+0\cdot x_3+0\cdot x_4+\cdots=1 \\ 0\cdot x_1+0\cdot x_2+9\cdot x_3+0\cdot x_4+\cdots=1 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{array}\right. \tag{4.5.1}$$

has the unique solution $x=\left\{(-1)^{k-1}\frac{1}{k^2}\right\}\in\mathfrak{l}$. This solution is neither finite nor directed.

4.6. The space $\mathfrak{c}$ is a subspace of the space $\mathfrak{l}^\infty$ of all bounded sequences $a=\{a_k\}$, $k\geqq 0$, with norm $\|a\|=\sup|a_k|$. The difficulty in dealing with $\mathfrak{l}^\infty$ lies in the intricate form of its continuous linear functionals. The situation here is quite different from that in $\mathfrak{c}$. *Thus the system* (0.1.1) *with the* $a_i=\{a_{ik}\}\in\mathfrak{l}^\infty$ *need have no minmal solutions in* $\mathfrak{l}$. Consider, for instance, the one equation

$$\sum_{k=0}^{\infty} a_k x_k=1, \tag{4.6.1}$$

where $0<a_k<1$ and $\sup a_k=1$. Clearly, for any solution $x=\{x_k\}\in\mathfrak{l}$ of (4.6.1) we must have $\|x\|=\sum_0^\infty |x_k|>1$. On the other hand, if $\xi_k^{(p)}=1/a_p$ when $k=p$, and $\xi_k^{(p)}=0$ otherwise then $\xi^{(p)}=\{\xi_k^{(p)}\}$ is a solution with $\|\xi^{(p)}\|=1/a_p$. Chosing suitable values of p, we see that $\inf\|x\|=1$, the inferior being taken with respect to all solutions x. It follows that (4.6.1) has no minimal solutions. *Suppose*, however, that the system (0.1.1), with $a_i\in\mathfrak{l}^\infty$, has minimal solutions $x^*\in\mathfrak{l}$. If then x^* and x^{**} are two minimal solutions, it is easy to see that

$$\arg x_k^*=\arg x_k^{**}, \tag{4.6.2}$$

whenever both $x_k^*\neq 0$ and $x_k^{**}\neq 0$, so that the *property* (β) (in the introduction) *for equal directedness still holds for the minimal solutions in this case, provided that they exist.*

For, clearly, $\frac{1}{2}(x^*+x^{**})$ is also a minimal solution, and hence

$$\sum_{k=0}^{\infty}\tfrac{1}{2}|x_k^*+x_k^{**}|=\tfrac{1}{2}\sum_{k=0}^{\infty}\left(|x_k^*|+|x_k^{**}|\right), \tag{4.6.3}$$

which implies $|x_k^*+x_k^{**}|=|x_k^*|+|x_k^{**}|$ for all k, or (4.6.2).

References.

[1] BANACH, S.: Théorie des opérations linéaires. Warszawa 1932. — [2] BOHNENBLUST, H. F., and A. SOBCZYK: Extensions of functionals on complex linear spaces. Bull. Amer. Math. Soc. **44**, 91—93 (1938). — [3] RIESZ, F.: Les systèmes d'équations linéaires à une infinité d'inconnues. Paris 1913. — [4] RIESZ, F., and B. SZ-NAGY: Leçons d'analyse fonctionelle. Budapest 1952. — [5] ROGOSINSKI, W. W., u. H. S. SHAPIRO: On certain extremum problems for analytic functions. Acta Math. **90**, 287—318 (1953).

King's College, University of Durham, Newcastle upon Tyne (England).

(Eingegangen am 7. Januar 1955.)

DINGHAS, A.
Math. Zeitschr. Bd. 63, S. 109—132 (1955)

Konvexitätseigenschaften von Mittelwerten harmonischer und verwandter Funktionen.

ISSAI SCHUR zum Gedächtnis.

Von
ALEXANDER DINGHAS.

§ 1. Einleitung.

1.1. Voraussetzungen. Im folgenden bedeutet H^n den Halbraum $x_1 > 0$ des euklidischen Raumes R^n, dargestellt durch die Kartesischen Koordinaten $x_1, \ldots, x_n$. Wir setzen ferner $x_1 = r\cos\vartheta$ mit $x_1^2 + \cdots + x_n^2 = r^2$ $(r > 0)$ und $-\pi < 2\vartheta < \pi$.

Die Funktionenklasse $L_0 = L_0(u)$ sei nun folgendermaßen definiert:

1. Die Funktion $u = u(P)$ ist in H^n eindeutig und genügt der (Lindelöfschen) Bedingung

$$\overline{\lim_{P \to Q}} u(P) \leq 0 \quad (P \in H^n) \tag{1.1.1}$$

für jeden endlichen Randpunkt Q von H^n.

2. In der Umgebung jedes Punktes $P \in H^n$, wo $u(P)$ endlich bleibt, ist sie entweder harmonisch oder eine zweimal stetig differenzierbare subharmonische Funktion. Das bedeutet, daß $u(P)$ in der Umgebung jedes Punktes $P \in H^n$, wo sie endlich bleibt, der Bedingung

$$\frac{\partial^2 u}{\partial x_1^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} \geq 0 \tag{1.1.2}$$

genügt.

Danach ist eine Funktion der Klasse L_0 stets eine zweimal differenzierbare subharmonische Funktion, welche der Lindelöfschen Bedingung (1.1.1) genügt.

Eine Erweiterung der Klasse L_0 und somit auch des Gültigkeitsbereiches der Ergebnisse dieser Arbeit erhält man durch Grenzprozesse. Im folgenden fassen wir alle subharmonischen Funktionen u, welche sich als Limes von nicht zunehmenden Folgen von subharmonischen Funktionen der Klasse L_0 darstellen lassen, in einer Klasse, der Klasse L_2 zusammen. Die Teilklasse von L_2, die monotone, gleichmäßig konvergente Folgen verwendet, soll mit L_1 bezeichnet werden.

Es gilt offenbar $L_0 \subseteqq L_1 \subseteqq L_2$. Ferner folgt aus der Definition der Klassen L_1 und L_2, daß Konvexitätssätze, die für die Klasse L_0 bewiesen werden, durch Grenzübergang ohne weiteres auf diese Klassen ausgedehnt werden können.

Es sei zunächst $u \in L_0$, und es bezeichne S_r die Halbkugel

$$x_1^2 + \cdots + x_n^2 = r^2 \quad (x_1 > 0,\ r > 0). \tag{1.1.3}$$

Setzt man

$$\psi(P) = \overset{+}{u}(P) : \cos\vartheta \tag{1.1.4}$$

mit

$$\overset{+}{u}(P) = \operatorname{Max}\{0, u(P)\}, \tag{1.1.5}$$

so wird sich zeigen, daß $\psi(P)$ auf S_r beschränkt ist.

Aus der allgemeinen Theorie der subharmonischen Funktionen weiß man jetzt, daß die Punktmenge $E\left(u(P) > 0,\ P \in H^n\right)$ (kurz E), falls sie nicht leer ausfällt, zusammenhängende Mengen enthält, die sich bis ins Unendliche erstrecken.

Man bezeichne jetzt mit $d\omega_n$ das Flächenelement der Einheitskugel S

$$\xi_1^2 + \cdots + \xi_n^2 = 1 \tag{1.1.6}$$

und bilde die Mittelwerte

$$m_k(r) = \int_{S_r} \psi^{2k}(P) \cos^2\vartheta \, d\omega_n. \tag{1.1.7}$$

Wird hier das Integral im Lebesgueschen Sinne genommen, so sind diese Funktionen für alle Funktionen der Klasse L_2 endlich und bei gegebenem u entweder für alle r gleich Null, oder von einem r an positiv.

Zweck der vorliegenden Arbeit ist zunächst, einige Konvexitätseigenschaften der Mittelwerte $m_k(r)$ zu studieren.

1.2. Formulierung der Hauptsätze. Wachstums- und Konvexitätseigenschaften der Mittelwerte (1.1.7) sind, soweit mir bekannt, bisher nur für $k = 1/2$ und $k = 1$ untersucht worden[1]). Der Fall eines beliebigen $k \geqq 1$ ist bisher in der Literatur noch nicht behandelt worden.

Die wichtigsten Ergebnisse der Arbeit lassen sich, was die Konvexität anbetrifft, durch folgende Sätze zum Ausdruck bringen:

Satz 1. *Für jedes $u \in L_2$ sind die Funktionen $m_k(r)^{1:k} r^{n-2}$ $(k \geqq 1)$ konvexe Funktionen von r^n.*

Satz 2. *Für jedes $u \in L_2$ sind die Funktionen $m_k(r)^{1:2k} : r$ konvexe Funktionen von $1 : r^n$.*

Will man nun den Grenzübergang $k \to \infty$ durchführen, so muß man zwei Fälle unterscheiden:

1. Die Funktion u liegt in der Klasse L_0 oder L_1. Dann gilt

$$\lim_{k \to \infty} m_k(r)^{1:2k} = \operatorname*{Max}_{P \in S_r} \{\psi(P)\} = M_1(r).$$

[1]) Man vgl. z.B. AHLFORS, L.: On Phragmén-Lindelöf's Principle. Trans. Amer. Math. Soc. **41**, 1—8 (1937). Hier werden ausschließlich harmonische Funktionen zugrunde gelegt, doch lassen sich die Ergebnisse ohne weiteres auf alle Funktionen der Klasse L_2 übertragen. — DINGHAS, A.: Über das Phragmén-Lindelöfsche Prinzip und einige andere verwandte Sätze. Sitzungsber. Preuß. Akad. Wissensch., Berlin, Math.-Physik. Kl. **12**, 123—140 (1938). — TSUJI, M.: On Harmonic Function on a Half-plane. Jap. J. of Math. **15**, 277—285 (1939).

2. Es ist $u \in L_2$. In diesem Falle definieren wir die Größe $M_1(r)$ folgendermaßen:

Es sei $P \in S_r$ und es bezeichne $M^*(r)$ das wahre Maximum von u^+ auf S_r, d.h. die untere Grenze aller Zahlen A, für welche die Menge $E(u(P) \geqq A, P \in S_r)$ auf S_r gemessen, den Lebesgueschen Inhalt Null hat. Dann wird $M_1(r) = M^*(r)$ gesetzt.

Durch diese Definition von $M_1(r)$ kann man nun folgende allgemeine Sätze beweisen:

Satz 3. *Die Funktion $M_1(r)^2 r^{n-2}$ ist eine konvexe Funktion von $1:r^n$.*

Aus dem zuletzt formulierten Satz 3 folgt ohne Schwierigkeit, daß, falls u nicht durchweg $\leqq 0$ ausfällt, der Ausdruck $m_k(r)^{1:2k}:r$, insbesondere auch der Ausdruck $M_1(r):r$ für $r \to 0$ gegen positive Zahlen, die auch unendlich sein können, konvergieren. Das würde bedeuten:

Für jedes $u \in L_1$ genügt nicht bloß u dem Maximumprinzip, sondern auch die Funktion $u(P):x_1$. Insbesondere trifft dies für den Fall zu, wenn u in H^n regulär harmonisch ist und der Lindelöfschen Bedingung (1.1.1) genügt.

Die Beweise findet der Leser in § 5.

Die beiden letzten Paragraphen der Arbeit beschäftigen sich mit folgender Frage: Kann man durch Erweiterung der Klasse erreichen, daß die Spanne bei der Aussage im klassischen Phragmén-Lindelöfschen Satz: ($u \leqq 0$ in H^n oder mindestens von der Ordnung Eins) verringert wird?

Es zeigt sich, daß diese Frage positiv beantwortet werden kann, und daß man durch Hinzuziehung einer Klasse superharmonischer Funktionen die vorhandene Lücke ausfüllen kann. Für die genaue Formulierung der Resultate möge auf die Entwicklungen von § 7 hingewiesen werden, besonders auf die allgemeine Ungleichung (7.2.7).

Der letzte Paragraph erhält allgemeine Konvexitätssätze über subharmonische und harmonische Funktionen in R^n, insbesondere eine Übertragung des Hadamardschen Dreikreisesatzes auf diese Klassen.

Das wichtigste Resultat kann hier folgendermaßen formuliert werden: Man definiere die Klassen L_0^*, L_1^*, L_2^* genau wie die Klassen L_0, L_1, L_2 mit dem Unterschied, daß an Stelle H^n der ganze Raum R^n genommen wird und die Lindelöfsche Bedingung (1.1.1) wegfällt. Ebenfalls definiere man $M(r)$ für die Klasse L_2^*, ähnlich wie $M_1(r)$ für die Klasse L_2. Setzt man dann

$$\mu_k(r) = \{\int u^+(P)^{2k} d\omega_n\}^{1:2k} \quad (k \geqq 1), \tag{1.1.8}$$

so gelten folgende Sätze:

Satz 4. *Die Funktion $\mu_k(r)$ $(k \geqq 1)$ ist eine konvexe Funktion von $1:r^{n-2}$.*

Satz 5. *Für $n=2$ ist $M(r)$ eine konvexe Funktion von $\log r$. Dagegen ist für $n>2$ $\mu_k(r)\, r^{n-2}$ eine konvexe Funktion von r^{n-2}.*

Geht man hier mit $k \to \infty$, so erhält man den

Satz 6. *Ist $n=2$, so ist $M(r)$ eine konvexe Funktion von $\log r$. Für $n>2$ ist $M(r)\, r^{n-2}$ eine konvexe Funktion von r^{n-2}.* Dieser Satz, der für $n=2$ mit

dem Hadamardschen Dreikreisesatz übereinstimmt, gibt für $n > 2$ seine natürliche Verallgemeinerung.

Die Bemerkungen des letzten Abschnitts von § 7 bewegen sich in der Richtung, die Frage der Existenz regulärer Integrale partieller Differentialgleichungen zu beantworten. Ich habe mich hier nur mit einigen Bemerkungen begnügt, da, soviel ich übersehe, die Arbeiten von Wittich diese Frage soweit gebracht haben, daß ich hier wesentlich neue Resultate dem Umfang nach kaum hinzufügen könnte.

1.2. Methode. Das Verfahren, das zum Beweis der hier formulierten Sätze verwendet wird, geht ursprünglich auf eine Methode von CARLEMAN[2]) zurück, der sie in etwas veränderter Form für $n = 2$ und $k = 1$ entwickelte. Der Leser wird leicht merken, wie stark diese Methode modifiziert werden muß, damit sie die allgemeinen und tieferliegenden Resultate der vorigen Nummern liefert.

Die gesamte Funktionentheorie einer komplexen Veränderlichen stützt sich eigentlich auf die Greensche Transformationsformel (4.1.3). Ich hoffe, in dieser Arbeit gezeigt zu haben, wie weit man auch Ungleichungsprobleme mit dieser Formel direkt angreifen und beantworten kann.

Im ganzen genommen ist diese Arbeit im Laufe eines Fortgeschrittenen-Kursus über Funktionentheorie, den ich in der Columbia University in New York mit Mitteln der Ford Foundation im Herbst 1952 gehalten habe, entstanden. Anderweitige Ablenkungen erlaubten mir jedoch erst jetzt[3]), das druckfertige Manuskript fertigzustellen.

Der Umweg über die Klasse L_0 und die Bildung der Klassen L_1 und L_2 scheint der natürlichste Weg zur Gewinnung der hier aufgezählten Resultate zu sein. Denn soweit es sich um subharmonische Funktionen handelt, scheinen wohl die Klasse L_1, L_2 sowie auch die Klasse L_1^*, L_2^* die allgemeinsten subharmonischen Funktionen zu enthalten. Ich hoffe, später bei Gelegenheit einer Darstellung dieses Fragenkomplexes (wie heutzutage üblich geworden ist) in einer kleinen Monographie dem Ganzen eine breitere Basis zu geben.

Zum Schluß möchte ich noch Herrn Dr. JOACHIM NITSCHE und Herrn KARL-HEINZ DIENER, für Hilfe bei der Korrekturarbeit und einige kritische Bemerkungen danken.

[2]) CARLEMAN, T.: Sur une inégalité differentielle dans la théorie des fonctions analytiques. Compt. Rend. Acad. Sci. Paris **196**, 995—997 (1933). Man vgl. auch DINGHAS, A.: Bemerkung zu einer Differentialgleichung von Carleman. Math. Z. **41**, 713—716 (1936). Ebenfalls: Sur un théorème de Denjoy-Carleman-Ahlfors. Bull. Soc. Math. Fr. **64**, 212 (1936).

[3]) Eine Zusammenfassung der Ergebnisse dieser Arbeit, allerdings nur eines Teiles, gab ich zuerst in einer Comptes Rendus Note: Sur quelques théorèmes concernant la convexité des moyennes d'une classe des fonctions sousharmoniques. Compt. Rend. Acad. Sci. Paris **237**, 594—595 (1953). — Für verwandte Probleme, zu welchen die Carlemansche Methode führt, vgl. man auch die Dissertation von HUBER: Über Wachstumseigenschaften gewisser Klassen von subharmonischen Funktionen. Zürich 1952. Dort findet der Leser auch weitere Literaturangaben. In ähnlicher Richtung bewegt sich die Dissertation von HANS-HEINRICH KELLER: Über das Anwachsen von Potentialfunktionen im dreidimensionalen Raum (Zürich 1950). Allgemeine Sätze und eine ausführliche Angabe der Eigenschaften der subharmonischen Funktionen gibt das ausgezeichnete Buch von T. RADO, Subharmonic Functions. Berlin 1937.

§ 2. Hilfssätze. Anfang.

2.1. Das Greensche Lemma und der Gauss-Ostrogradskische Satz. Damit der Leser alles beisammen haben möge, geben wir hier noch kurz den Wortlaut der beiden Sätze, auf den sich die nachfolgenden Entwicklungen stützen. Dabei werden soviel Voraussetzungen gemacht, als es für die Anwendung notwendig ist.

Wir bezeichnen im folgenden als regulären Bereich, eine beschränkte Punktmenge von R^n, die sich als Summe von endlich vielen Körpern im Sinne von ERHARD SCHMIDT[4]) darstellen läßt. Damit der Leser nicht die ganze Theorie zu lesen braucht, soll hier nur erwähnt werden, daß der Durchschnitt H der Punktmenge $E\left(u \geqq \frac{1}{m}\right)$ $(m = 1, 2, 3, \ldots)$ mit jeder Halbkugel $x_1^2 + \cdots + x_n^2 \leqq r^2$ $(x_1 > 0)$ von H^n ein regulärer Bereich ist, sobald $u \in L_0$ gilt.

Es bezeichne nun G den offenen Kern eines regulären Bereiches und Γ seinen Rand. Wir setzen dann $\overline{G} = G + \Gamma$. Ferner bezeichne ν die Richtung der äußeren Normale von $\overline{G}$, $d\tau$ das Volumenelement von G und dO das Oberflächenelement von Γ. Ist $\Phi = \Phi(P)$ eine in $\overline{G}$ stetig differenzierbare Funktion der Koordinaten $x_1, \ldots, x_n$ von P, so gilt für jedes k $(k = 1, \ldots, n)$

$$\int_{\overline{G}} \frac{\partial \Phi}{\partial x_k} d\tau = \int_{\Gamma} \Phi \cos(\nu, x_k)\, dO. \tag{2.1.1}$$

Diese Formel drückt das klassische Greensche Lemma aus. Hat man n Funktionen Φ_k mit den erwähnten Eigenschaften, so erhält man durch Summation

$$\int_{\overline{G}} \left\{ \sum_1^n \frac{\partial \Phi_k}{\partial x_k} \right\} d\tau = \int_{\Gamma} (\Phi, \nu)\, dO, \tag{2.1.2}$$

wobei hier Φ den Vektor mit den Koordinaten Φ_k bedeutet und die Klammer das innere Produkt von Φ mit ν darstellt.

Diese Gleichung drückt den Gauß-Ostrogradskischen Satz aus.

Der Fall, der uns im folgenden interessieren wird, wird durch die Gleichungen $\Phi_k = v \frac{\partial u}{\partial x_k}$ gegeben, wobei u, v zweimal stetig differenzierbare Funktionen bedeuten.

2.2. Vorbereitende Tatsachen. Einem Gebiet T von R^n sei die Metrik

$$ds^2 = g_{ik}\, dx^i\, dx^k \quad [5]) \tag{2.2.1}$$

aufgeprägt.

Wir betrachten einen regulären Bereich $\overline{G} \subseteq T$ und setzen voraus, daß wir eine Funktion $v = v(P)$ mit folgenden Eigenschaften bestimmen können:

1. Die Funktion $v(P)$ ist zweimal stetig differenzierbar in G und stetig in $G + \Gamma$.

[4]) SCHMIDT, ERHARD: Über das isoperimetrische Problem im Raum von n Dimensionen. Math. Z. **44**, 689—788 (1939).

[5]) Doppelte Indizes ziehen, wie üblich, eine Summation von 1 bis n nach sich, unabhängig voneinander.

2. Es ist $v(P)>0$ für $P\in G$.

3. Die Funktion $v(P)$ genügt in G der Differentialgleichung.

$$\Delta v+\lambda_0 v=0 \quad [6] \tag{2.2.2}$$

mit einem $\lambda_0>0$. Dabei bedeutet allgemein Δu den Lamé-Beltramischen Differentialausdruck

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial x^i}\left(\sqrt{g}\,g^{ik}\frac{\partial u}{\partial x^k}\right) \quad . \quad (g=|g_{ik}|>0). \tag{2.2.3}$$

Es bezeichne nun U die Klasse der Funktionen u mit folgenden Eigenschaften:

1. Jedes u ist in G zweimal stetig differenzierbar, stetig in $G+\Gamma$, und Null auf Γ.

2. Setzt man

$$\varphi(P)=u(P):v(P), \tag{2.2.4}$$

so gilt für jeden Randpunkt $Q\in\Gamma$

$$\overline{\lim_{P\to Q}}\,|\varphi(P)|<+\infty. \tag{2.2.5}$$

3. Die Größen $\left|\frac{\partial u}{\partial x^k}\right|$ $(k=1, 2, \ldots, n)$ sind gleichmäßig beschränkt in G.

Definiert man, wie üblich, $\nabla(u, v)$ durch die Gleichung

$$\nabla(u,v)=g^{ik}\frac{\partial u}{\partial x^i}\frac{\partial v}{\partial x^k}, \tag{2.2.6}$$

so gilt für jedes $u\in U$

$$\int_G \nabla(u,u)\,d\tau_n<+\infty. \tag{2.2.7}$$

Dabei bedeutet $d\tau_n$ das Volumenelement $\sqrt{g}\,dx^1\ldots dx^n$ von G. Das Integral links wird hier (wie auch im folgenden) als Limes von Integralen $\int_{\bar{G}_k}\nabla(u,u)\,d\tau_n$ mit $\bar{G}_k\leqq\bar{G}_{k+1}<G$ $(k=1, 2, \ldots)$ und $\lim_{k\to\infty}\bar{G}_k=G$ definiert.

Sind $u_1, u_2\in U$, so gilt nach der Greenschen Transformation

$$\int_G u_1\Delta u_2\,d\tau_n=-\int_G \nabla(u_1,u_2)\,d\tau_n. \tag{2.2.8}$$

Infolge der Ungleichung

$$\left(\int_G|\nabla(u_1,u_2)|\,d\tau_n\right)^2\leqq\int_G\nabla(u_1,u_1)\,d\tau_n\int_G\nabla(u_2,u_2)\,d\tau_n,$$

sind beide Seiten dieser Gleichung endlich.

2.3. Der Haupthilfssatz. Die Verallgemeinerung der Carlemanschen Methode wird nun durch folgenden wichtigen Hilfssatz ermöglicht:

Hilfssatz 1. *Die Funktion $f(x)$ möge im Intervall $I-\infty<x<+\infty$ folgende Eigenschaften besitzen:*

[6] Für die allgemeinen Begriffsbildungen vgl. man: COURANT u. HILBERT: Methoden der mathematischen Physik, Bd. 1. Berlin 1931.

1. In jedem Punkt $x \in I$ besitzt $f(x)$ eine stetige Ableitung $f'(x) > 0$.
2. Es ist $f(0) = 0$.
Dann gilt für jedes $u \in U$

$$\int_G f(\varphi)\, v\,(\Delta u + \lambda_0 u)\, d\tau_n \leqq 0, \tag{2.3.1}$$

und das Gleichheitszeichen gilt dann und nur dann, wenn in G $u = Cv$ mit einem konstanten C gilt, und $v(P)$ auf Γ verschwindet.

Beweis. Es sei $\overline{G}' \subset G$. Man betrachte das Integral

$$J = \int_{\overline{G}'} v^2 V\left(\frac{u}{v}, \frac{u}{v}\right) d\tau_n, \tag{2.3.2}$$

wegen $v > 0$ auf dem Rand Γ' von $\overline{G}'$ kann man den Integranden entwickeln und schreiben

$$J = \int V(u,u)\, d\tau_n - 2\int \frac{u}{v} V(u,v)\, d\tau_n + \int \frac{u^2}{v^2} V(v,v)\, d\tau_n.$$

die Integration ist hier über $\overline{G}'$ zu erstrecken.

Es bedeuten nun ν und dO die äußere Normale bzw. das Oberflächenelement von Γ'. Dann wird durch Anwendung des Greenschen Lemmas

$$-2\int \frac{u}{v} V(u,v)\, d\tau_n = \int u^2 \frac{\partial}{\partial x^i}\left(\sqrt{g}\, g^{ik} \frac{1}{v} \frac{\partial v}{\partial x^k}\right) d\tau_n + J_0,$$

wobei J_0 das Flächenintegral $-\int_{\Gamma'} \frac{u^2}{v} \frac{\partial v}{\partial \nu}\, dO$ darstellt. Schreibt man jetzt J_1 für das Integral rechts, so wird nach einfachen Rechnungen

$$J_1 = \int \frac{u^2}{v} V v\, d\tau_n - \int \frac{u^2}{v^2} \Delta(v,v)\, d\tau_n.$$

Faßt man diese Ergebnisse zusammen, so erhält man mit Rücksicht auf (2.2.2)

$$J = J_0 + \int_{\overline{G}'} \{V(u,u) - \lambda_0 u^2\}\, d\tau_n.$$

Man schöpfe nun G durch eine monotone Folge $(\overline{G}'_k)$ $(k = 1, 2, \ldots)$ aus und beachte, daß bei genügend großem k der Betrag von u auf $\Gamma'_k \leqq \varepsilon$ ausfällt. Läßt man also $k \to \infty$ konvergieren, so konvergieren die entsprechenden Integrale J_0 gegen Null, und man erhält

$$\int_G \{V(u,u) - \lambda_0 u^2\}\, d\tau_n = \int_G v^2 V\left(\frac{u}{v}, \frac{u}{v}\right) d\tau_n. \tag{2.3.3}$$

Es seien nun $u_1, u_2 \in U$. Man ersetze in dieser Formel u durch die lineare Kombination $\alpha_1 u_1 + \alpha_2 u_2$ und vergleiche auf beiden Seiten die Koeffizienten von $\alpha_1 \alpha_2$.

Man erhält dann

$$\int_G \{V(u_1, u_2) - \lambda_0 u_1 u_2\}\, d\tau_n = \int_G v^2 V\left(\frac{u_1}{v}, \frac{u_2}{v}\right) d\tau_n. \tag{2.3.4}$$

Jetzt setze man hier

$$u_1 = u \qquad (u \in U), \qquad u_2 = f(\varphi)\, v \tag{2.3.5}$$

und wende das Greensche Lemma in der Form

$$\int_G V(u_1, u_2)\, d\tau_n = -\int_G u_2\, \Delta u_1\, d\tau_n \tag{2.3.6}$$

an. Dann wird

$$\int_G f(\varphi)\, v(\Delta u + \lambda_0 u)\, d\tau_n = -\int_G f'(\varphi)\, v^2\, V\left(\frac{u}{v}, \frac{u}{v}\right) d\tau_n, \tag{2.3.7}$$

womit auch der Hilfssatz bewiesen wurde. Ich möchte hier erwähnen, daß dieser Hilfssatz auch unter denjenigen schwächeren Voraussetzungen gilt, unter welchen bei Zugrundelegung des Lebesgueschen Integrals die Greenschen Umformungen gelten. Ich gehe jedoch darauf nicht weiter ein.

§ 3. Hilfssätze. Ende.

3.1. Spezialfälle. Im folgenden ist der Fall, wo $f(x) = x^{2k-1}$ $(k \geqq 1)$ ist und T mit der Halbkugel

$$\xi_1^2 + \cdots + \xi_n^2 = 1 \qquad (\xi_1 = \cos\vartheta \geqq 0) \tag{3.1.1}$$

zusammenfällt, von besonderem Interesse. Wir zeigen vorerst, daß für jede Funktion $u \in L_0$ der Quotient $\overset{+}{u}(P) : \cos\vartheta$ auf jedem S_r beschränkt bleibt.

Hilfssatz 2. *Die Funktion*

$$\psi(P) = \overset{+}{u}(P) : \cos\vartheta \qquad (P \in S_r,\ u \in L_0), \tag{3.1.2}$$

ist in jedem Intervall $0 < r_1 \leqq r \leqq r_2 < +\infty$ *nach oben beschränkt.*

Beweis. Ist $P(x_1, \ldots, x_n) \in H^n$, so soll P' den Punkt von R^n mit den Koordinaten $-x_1, x_2, \ldots, x_n$ bezeichnen. Es sei ferner H_R die Halbkugel

$$x_1^2 + \cdots + x_n^2 < R^2 \qquad (x_1 > 0) \tag{3.1.3}$$

und $P \in H_R$. Wir setzen $OP = r$, $\varrho(Q) = PQ$ und $\varrho'(Q) = P'Q$. Ferner soll Q' den Punkt der Strecke OQ mit $OQ' = 1$ angeben.

Die Funktion

$$V(P) = \frac{R^{n-2}}{\omega_n} \int_{S_R} \overset{+}{u}(Q)\, \{\varrho^{-n}(Q) - \varrho'^{-n}(Q)\}\, (R^2 - r^2)\, d\omega_n(Q'),$$

wobei ω_n die Gesamtoberfläche von S bedeutet, ist in H_R harmonisch[7]), nimmt auf S_R die Randwerte $\overset{+}{u}(Q)$ und verschwindet auf dem ebenen Randteil $x_1 = 0$ von H_R. Da $u(P)$ subharmonisch und $u \in L_0$ ist, so gilt in H_R

$$u(P) \leqq V(P). \tag{3.1.4}$$

Nun ist, wenn $y_1, \ldots, y_n$ die Koordinaten von $Q \in S_R$ bezeichnen

$$0 \leqq \varrho^{-n}(Q) - \varrho'^{-n}(Q) \leqq 2n\, x_1 y_1 : (R - r)^{n+1},$$

[7]) Man vgl. z. B. CARATHÉODORY, C.: On Dirichlets Problem. Amer. J. Math. **59**, 709 (1937). Dieser Hinweis gilt natürlich nur für die dort entwickelte Poissonsche Formel. Der Leser kann daraus (3.1.4) selbst ableiten.

und infolgedessen wird für $P \in S_r$, $R = 2r$

$$\overset{+}{u}(P) : \cos\vartheta \leqq C(r), \tag{3.1.5}$$

wobei $C(r)$ endlich ist. Das beweist den Hilfssatz. Es sei nun $u \in L_0$ und $E(u > 0)$ nicht leer. Bei genügend großem m ist dann $E\left(u \geqq \frac{1}{m}\right)$ ebenfalls nicht leer.

Schreibt man hier wieder u an Stelle $u - \frac{1}{m}$, so gilt der

Hilfssatz 3. *Für jedes $k \geqq 1$ gilt*

$$\int_{E_r} \psi^{2k-2}(P)\,(\Delta u + (n-1)\,u)\,u\,d\omega_n \leqq 0, \tag{3.1.6}$$

und das Gleichheitszeichen tritt nur dann ein, wenn in H^n $u = C x_1$ mit einem konstanten C ist.

Hier bedeutet Δu den Beltramischen Operator (2.2.3) geschrieben mit Hilfe der g_{ik} der Einheitskugel $\xi_1^2 + \cdots + \xi_n^2 = 1$ und E_r die Menge $E(u \geqq 0) \cap S_r$.

Bei der Anwendung des Hilfssatzes 1 hat man zuerst zu berücksichtigen, daß $\cos\vartheta$ der Differentialgleichung (2.2.2) mit $\lambda_0 = n - 1$ genügt[8]), und daß in jedem Punkt von E_r $\cos\vartheta > 0$ ausfällt. Ferner ist zu berücksichtigen, daß $\psi^{2k-2}(P)\cos\vartheta$ $(u \in L_0)$ ebenfalls der Klasse U angehört, da ja $\psi^{2k-1}\cos\vartheta = u\,\psi^{2k-2}$ ist, und der Exponent nicht negativ ausfällt.

Der Hilfssatz 3 gilt selbstverständlich für viel breitere Funktionenklassen, als die Klasse L_0, z.B. für die Klasse L_2. Er wird aber ausschließlich für diese Klasse gebraucht und ich halte es nicht für notwendig, auf Vertiefungen einzugehen. Lediglich für den Fall $k = 1$ möchte ich hier eine Bemerkung anschließen, und zwar um den Eindruck zu zerstreuen, daß die hier gemachten Voraussetzungen enger sind als diejenigen des klassischen Wirtingerschen Lemmas[9]) für den Fall $n = 2$.

Wir setzen voraus, u sei auf der abgeschlossenen Halbkugel S_1

$$\xi_1^2 + \cdots + \xi_n^2 = 1 \qquad (\xi_1 \geqq 0), \tag{3.1.7}$$

als eindeutige Funktion von $P(\xi_1, \ldots, \xi_n)$ definiert und setzen

$$\xi_k = \sin\varphi_k\,\Pi_{k-1} \quad (k = 1, 2, \ldots, n-1), \qquad \xi_n = \Pi_n, \tag{3.1.8}$$

mit $\Pi_0 = 1$ und $\Pi_k = \prod_1^k \cos\varphi_r$. Die Variationsintervalle der φ_r sind hier einleuchtend.

Wir setzen nun ferner voraus:

1. Die Funktion u ist in S_1 stetig und genügt der Bedingung $u(0, \xi_2, \ldots, \xi_n) = 0$.

2. Die Funktion u ist in φ_1 totalstetig in der Umgebung von $\xi_1 = 0$.

3. Formt man u durch (3.1.8) um, so existiert $V(u, u)$ fast überall und ist im Lebesgueschen Sinne integrierbar. Dann kann bewiesen werden:

---

[8]) Vgl. z.B. Dinghas, A.: Zur Theorie der konvexen Körper im n-dimensionalen Raum. Abh. Akad. Wiss. Berlin **1939**, No. 4.

[9]) Vgl. z.B. Hardy-Polya-Littlewood: Inequalities. Cambridge 1934.

Es gilt stets

$$\int_{S_1} \{V(u,u) - (n-1)\,u^2\}\,d\omega_n \geqq 0, \tag{3.1.9}$$

und das Gleichheitszeichen tritt nur für den Fall $u = C \sin \varphi_1$ ein.

Es bedeute in der Tat S_ϑ die Kugelzone $\sin \varphi_1 \leqq \sin \vartheta$ und es bezeichne Γ_ϑ den Rand $\varphi_1 = \vartheta$ von S_ϑ.

Dann gilt, wie man leicht sieht,

$$\int_{\Gamma_\vartheta} \frac{u^2}{\sin \vartheta}\,dO_{n-1} \leqq K \int_{S_\vartheta} V(u,u)\,d\omega_n.$$

Da nun wegen 3. die rechte Seite ein $o(1)$ ist, so wird

$$\int_{S_1} \{V(u,u) - (n-1)\,u^2\}\,d\omega_n = \int_{S_1} v^2\,V\left(\frac{u}{v}, \frac{u}{v}\right) d\omega_n,$$

mit $v = \sin \varphi_1$. Das beweist (3.1.9).

§ 4. Verallgemeinerung der Carlemanschen Methode.

4.1. Eine allgemeine Differentialungleichung. Vom Standpunkt der Greenschen Transformation aus gesehen erscheint die Carlemansche Methode als direkte Anwendung einer allgemeinen Transformationsformel, die in Spezialfällen mehrfache Anwendung findet. Es sei G ein regulärer Bereich und es seien dort $U = U(x_1, \ldots, x_n)$, $V = V(x_1, \ldots, x_n)$ zwei zweimal stetig differenzierbare Funktionen der rechtwinkligen Koordinaten $x_1, \ldots, x_n$. Man setze

$$(\operatorname{grad} U, \operatorname{grad} V) = \sum_1^n \frac{\partial U}{\partial x_k} \frac{\partial V}{\partial x_k} \tag{4.1.1}$$

und

$$\operatorname{div} \operatorname{grad} U = \sum_1^n \frac{\partial^2 U}{\partial x_k^2}. \tag{4.1.2}$$

Dann gilt nach dem in 2.1 entwickelten Gaußschen Satz

$$\int_\Gamma V \frac{\partial U}{\partial \nu}\,dO_{n-1} = \int_{G+\Gamma} \{V \operatorname{div} \operatorname{grad} U + (\operatorname{grad} U, \operatorname{grad} V)\}\,d\tau_n. \tag{4.1.3}$$

Hierbei bedeutet $d\tau_n$ das Volumenelement $dx_1 \ldots dx_n$ von G und dO_{n-1} das Flächenelement des Randes Γ von G.

Man betrachte nun eine Funktion der Klasse L_0 und setze voraus, daß es mindestens einen Punkt P von H^n gibt, in dem $u(P) > 0$ ist. Betrachtet man dann die Funktionen $u_m = u - \frac{1}{m}$ $(m = 1, 2, \ldots)$, so gehören diese ebenfalls der Klasse L_0 an und genügen von einem m_0 an der Bedingung $u_m(P) > 0$. Man wähle nun ein festes $m > m_0$, schreibe der Einfachheit halber u an Stelle

von u_m und wende (4.1.3) auf den Durchschnitt G_r von $E(u>0)$ mit der Halbkugel S_r

$$x_1^2+\cdots+x_r^2\leqq r^2 \qquad (x_1\geqq 0) \tag{4.1.4}$$

an, indem man

$$v=2k\,\psi^{2k-1}(P)\cos\vartheta \qquad (\psi=\overset{+}{u}:\cos\vartheta)$$

wählt.

Nun wird

$$(\operatorname{grad}u,\operatorname{grad}v)=u_r v_r+\tfrac{1}{r^2}V(u,v), \tag{4.1.5}$$

und

$$\operatorname{div}\operatorname{grad}u=\tfrac{1}{r^{n-1}}(r^{n-1}u_r)_r+\tfrac{1}{r^2}\Delta u, \tag{4.1.6}$$

wobei $V(u,v)$, Δu die Beltramischen Differentialausdrücke von 3.1 für die Einheitskugel (3.1.1) bedeuten. Beachtet man jetzt, daß v auf dem Randteil von Γ, der nicht auf S_r liegt, verschwindet, so erhält man mit $v=\cos\vartheta$

$$r^{n-1}\int\limits_{E_r} v\frac{\partial u}{\partial r}\,d\omega_n=\int\limits_{G_r} f(P)\,t^{n-1}\,dt\,d\omega_n, \tag{4.1.7}$$

wobei der Kürze wegen

$$f(P)=v\left\{\tfrac{1}{r^{n-1}}(r^{n-1}u_r)_r+\tfrac{1}{r^2}\Delta u\right\}+\left\{u_r v_r+\tfrac{1}{r^2}V(u,v)\right\}$$

gesetzt wurde.

Schreibt man allgemein

$$F(x)=\int\limits_{E_x} f(P)\,d\omega_n, \tag{4.1.8}$$

so erhält man

$$r^{n-1}\int\limits_{E_r} v\frac{\partial u}{\partial r}\,d\omega_n=\int\limits_0^r x^{n-1}F(x)\,dx. \tag{4.1.9}$$

Will man nun diese Gleichung differenzieren, so sieht man leicht ein, daß stets

$$\frac{d}{dr}\left\{r^{n-1}\int\limits_{E_r} v\frac{\partial u}{\partial r}\,d\omega_n\right\}=r^{n-1}F(r) \tag{4.1.10}$$

ist, und zwar für diejenigen r, für welche $\int d\omega_n$ erstreckt über die Menge $E\left(u(P)=0,\ P\in S_r\right)$ verschwindet. In allen übrigen Stellen ist die untere Derivierte der linken Seite von (4.1.5) nicht kleiner als $r^{n-1}\int f(P)\,d\omega_n$, wobei das Integral über diejenige Teilmenge Γ_r von S_r zu erstrecken ist, auf der $u(P)>0$ ist.

Setzt man jetzt

$$m_k=m_k(r)=\int\limits_{\Gamma_r}\psi^{2k}(P)\cos^2\vartheta\,d\omega_n \qquad (k\geqq 1), \tag{4.1.11}$$

so ist zunächst

$$m_k' = 2k \int_{\Gamma_r} \psi^{2k-1}(P)\, u_r \cos\vartheta \, d\omega_n,$$

und

$$\int_{\Gamma_r} u_r v_r \, d\omega_n = 2k(2k-1) \int_{\Gamma_r} \psi^{2k-2}(P)\, u_r^2 \, d\omega_n. \tag{4.1.12}$$

Versteht man also unter $(r^{n-1} m_k')'$ in den Stellen, wo diese Ableitung nicht existiert, die untere Derivierte, und beachtet man, daß u subharmonisch ist, so wird

$$\frac{1}{r^{n-1}} (r^{n-1} m_k')' \geqq 2k(2k-1) \int_{\Gamma_r} \psi^{2k-2}\, u_r^2 \, d\omega_n + \frac{1}{r^2} \int_{\Gamma_r} V(u, v)\, d\omega_n.$$

Mit Hilfe der Greenschen Transformation erhält man

$$-\int_{\Gamma_r} V(u, v)\, d\omega_n = \int_{\Gamma_r} v\, \Delta u \, d\omega_n = 2k \int_{\Gamma_r} \psi^{2k-1}(P)\, \Delta u \cos\vartheta \, d\omega_n.$$

Ferner gilt

$$m_k'^2 \leqq 4k^2 m_k \int_{\Gamma_r} \psi^{2k-2}(P)\, u_r^2 \, d\omega_n.$$

Faßt man alle diese Resultate zusammen, so erhält man

$$\frac{1}{r^{n-1}} (r^{n-1} m_k')' \geq \frac{2k-1}{2k} \frac{m_k'^2}{m_k} + \frac{2k(n-1)}{r^2} m_k. \tag{4.1.13}$$

Es darf hier wiederholt werden, daß u stets $u - \frac{1}{m}$ bedeutet, und daß unter Umständen links die untere Derivierte von $r^{n-1} m_k'$ gemeint ist. Wegen der Stetigkeit von m_k' ist dies jedoch für das gesteckte Ziel unerheblich.

Die Ungleichung (4.1.13) ist für die nachfolgenden Ausführungen grundlegend. Nimmt man $k=1$, $n=2$ an, so wird

$$\frac{1}{r} (r\, m_1')' \geq \frac{1}{2} \frac{m_1'^2}{m_1} + \frac{2m_1}{r^2}, \tag{4.1.14}$$

d. h. die zu dem hier betrachteten Fall der Halbebene Carlemansche Ungleichung.

4.2. Umformungen. Anfang. Ist die Funktion $u\,(u \in L_0)$ in H^n nicht durchweg ≤ 0, so ist von einem r an $m_k > 0$. Wir schreiben (4.1.13) in der Form

$$\frac{m_k''}{m_k} + \frac{n-1}{r} \frac{m_k'}{m_k} - \frac{2k-1}{2k} \left(\frac{m_k'}{m_k}\right)^2 \geq \frac{2k(n-1)}{r^2}$$

und formen diese Ungleichung folgendermaßen um.

Man setze zunächst

$$\mu(r) = m_k^{1:k}. \tag{4.2.1}$$

Dann erhält man für μ die Differentialungleichung

$$2\left(\frac{\mu''}{\mu} + \frac{n-1}{r} \frac{\mu'}{\mu}\right) - \left(\frac{\mu'}{\mu}\right)^2 \geq \frac{4(n-1)}{r^2}. \tag{4.2.2}$$

Führt man noch die Veränderliche ξ durch die Gleichung $\xi = r^n$ ein, und setzt man noch

$$y(\xi) = \mu(r)\, r^{n-2} = m_k(r)^{1/k}\, r^{n-2}, \tag{4.2.3}$$

so wird, falls die Ableitungen von y nach ξ durch $\dot{y}, \ddot{y}, \ldots$ bezeichnet werden,

$$\frac{\mu'}{\mu} = n\,\frac{\dot{y}}{y}\, r^{n-1} - \frac{n-2}{r}$$

und

$$\mu' = n\,\dot{y}\, r - \frac{n-2}{r^{n-1}}\, y.$$

Daraus folgt

$$\mu'' = n^2\,\ddot{y}\, r^n - n(n-3)\,\dot{y} + \frac{(n-1)(n-2)}{r^n}\, y,$$

und somitwird

$$\frac{\mu''}{\mu} + \frac{n-1}{r}\,\frac{\mu'}{\mu} = n^2\, r^{2n-2}\,\frac{\ddot{y}}{y} + 2n\, r^{n-2}\,\frac{\dot{y}}{y}.$$

Durch Einsetzen in (4.2.2) erhält man dann für y die Differentialungleichung

$$\ddot{y} \geqq \frac{1}{2}\, y \left(\frac{\dot{y}}{y} - \frac{1}{\xi}\right)^2,$$

woraus dann schließlich

$$\frac{d^2 y}{d\xi^2} \geqq \frac{1}{2}\, y \left(\frac{d}{d\xi} \log \frac{y}{\xi}\right)^2 \tag{4.2.1}$$

folgt.

4.3. Umformungen. Ende. Eine letzte Umformung von (4.1.13) sei noch gegeben. Man setze $\mu = g^2$ und $q = g : r = \mu^{\frac{1}{2}} : r = m_k^{1:2k} : r$.

Dann liefert diese Substitution die Differentialungleichungen

$$\frac{g''}{g} + \frac{n-1}{r}\,\frac{g'}{g} \geqq \frac{n-1}{r^2} \tag{4.3.1}$$

und

$$r\, q'' + (n+1)\, q' \geqq 0. \tag{4.3.2}$$

Da diese letztere sich in der Form

$$(r^{n+1}\, q')' \geqq 0 \tag{4.3.3}$$

schreiben läßt, so genügt q als Funktion von $\eta = 1 : r^n$ der Differentialungleichung

$$\frac{d^2 q}{d\eta^2} \geqq 0. \tag{4.3.4}$$

Hiermit sind alle Hilfsmittel bereitgestellt worden, die zum Beweis der in der Einleitung formulierten Sätze benötigt werden.

§ 5. Anwendungen.

5.1. Beweis der Sätze 1 bis 3. Da Konvexitätseigenschaften durch Anwendung von Grenzprozessen erhalten bleiben, so genügt es, die zwei ersten Sätze von 1.2 zunächst für Funktionen von der Form $u - \frac{1}{m}$ $(m = 1, 2, \ldots)$

mit $u \in L_0$ zu beweisen. Läßt man dann $m \to \infty$ konvergieren, so sind zunächst die Sätze 1 und 2 für jedes $u \in L_0$ bewiesen. Ein weiterer Grenzübergang $k \to \infty$ bei festem $u \in L_0$ beweist den Satz 3 für die Klasse L_0.

Damit der Leser alles beisammen haben möge, soll hier der bekannte Schluß wiederholt werden:

Es sei $f(x)$ in $a \leq x \leq b$ eindeutig stetig und ≥ 0. Dann gilt

$$J = \lim_{k \to \infty} \left\{ \int_a^b f(x)^k \, dx \right\}^{1:k} = \operatorname*{Max}_{a \leq x \leq b} \{f(x)\} = M. \tag{5.1.1}$$

In der Tat gilt zunächst trivialerweise $J \leq M$. Ist nun $\varepsilon > 0$, so gibt es ein Intervall $J_\varepsilon \langle a, b \rangle$ von der Länge $c J_\varepsilon > 0$, auf dem $f(x) \geq M - \varepsilon$ gilt. Daraus folgt $J \geq \lim_{k \to \infty} (c J_\varepsilon)^{1:k} (M - \varepsilon)$ also $J \geq M - \varepsilon$.

Wie verändert sich dieser Satz, wenn man die Stetigkeit aufhebt und dafür lediglich Meßbarkeit verlangt?

Wir zeigen:

Es sei M_0 das wahre Maximum von $f(x)$, d.h. die untere Grenze der Zahlen A, für welche die Menge $E(f(x) \geq A)$ das Lebesgue-Maß Null hat.

Dann gilt

$$\lim_{k \to \infty} \left\{ \int_a^b f(x)^k \, dx \right\}^{1:k} = M_0. \tag{5.1.2}$$

Es sei $\varepsilon > 0$. Man setze $f_\varepsilon(x) = f(x)$ für alle x, wo $f(x) \leq M_0 + \varepsilon$ ist und sonst $f_\varepsilon(x) = 0$.

Dann ist wegen

$$\int_a^b f^k(x) \, dx = \int_a^b f_\varepsilon^k(x) \, dx,$$

zunächst $J \leq M_0 + \varepsilon$. Andererseits ist

$$\int_a^b f^k(x) \, dx \geq (M_0 - \varepsilon)^k m_\varepsilon$$

mit einem $m_\varepsilon > 0$. Daraus folgt $J \geq M_0 - \varepsilon$, d.h. die Behauptung.

Ist nun $u \in L_0$, so kann ohne weiteres als bewiesen angesehen werden, daß $m_k(r) \, r^{n-2}$ als Grenzwert von monoton wachsenden konvexen Funktionen, d.h. der mit Hilfe der Funktionen $u - \frac{1}{m}$ gebildeten Ausdrücke (4.1.11), ebenfalls konvex ist. Darunter wird allgemein die Ungleichung

$$f(x) \leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2) \tag{5.1.1}$$

für jedes Wertetripel $x_1 < x < x_2$ verstanden.

Ist $u \in L_2$, so gibt es eine monoton abnehmende Folge (u_m) $(m = 1, 2, \ldots)$ von zweimal stetig differenzierbaren subharmonischen Funktionen der Klasse

L_0 mit $\lim u_m(P) = u(P)$ $(P \in H^n)$. Faßt man dann die Ausdrücke (4.1.11) als Lebesguesche Integrale auf, so erhält man durch Grenzübergang $m \to \infty$ den Beweis der Sätze 1 und 2 für die Klasse L_1. Faßt man schließlich für jedes $u \in L_2$ den Ausdruck $M_1(r)$ im Sinne der Gl. (5.1.2) auf, so genügt der Grenzübergang $k \to \infty$, um auch den Satz 3 für die Klasse L_2 zu beweisen.

Die Eigenschaft der Konvexität derAusdrücke $m_k^{1:k}\, r^{n-2}$ bzw. $m_k^{1:2k} : r$ als Funktion von r^n bzw. $1 : r^n$ zieht die Existenz der Grenzwerte

$$a_k = \lim_{r \to \infty} m_k^{1:2k}(r) : r \tag{5.1.2}$$

und

$$b_k = \lim_{r \to 0} m_k^{1:2k}(r)\, r^{n-1} \tag{5.1.3}$$

nach sich.

Insbesondere existiert Grenzwert

$$a = \lim_{r \to \infty} M_1(r) : r. \tag{5.1.4}$$

Wie nochmals darauf hingewiesen werden darf, ist $M_1(r)$ nicht durch die Maximumeigenschaft definiert, sondern durch die Bedingung, daß die Teilmenge von S_r für die $u \geqq M_1(r) + \varepsilon$ ist, das Maß Null (auf S_r!) hat, während dies für $M_1(r) - \varepsilon$ nicht zutrifft.

5.2. Scharfe Fassung des Phragmén-Lindelöfschen Prinzips. Ist $u \in L_1$, so kann man das klassische Phragmén-Lindelöfsche Prinzip schärfer als bisher fassen.

Es bedeute wieder H_r die Halbkugel

$$x_1^2 + \cdots + x_n^2 < r^2 \qquad (x_1 > 0,\ r > 0) \tag{5.2.1}$$

und es sei $P(x_1, \ldots, x_n) \in H_r$.

Dann kann man beweisen:

Entweder ist die Funktion

$$m(P) = \overset{+}{u}(P) : x \tag{5.2.2}$$

in H_r identisch Null oder sie ist stetig und nimmt ihr Maximum auf dem Rande S_r von H_r an.

Der Beweis des nichttrivialen Teiles dieses Satzes folgt aus der Tatsache, daß für je zwei Punkte $P(x_k) \in S_r$ und $P'(x_k') \in S_{r'}$ $(r < r')$

$$\frac{\overset{+}{u}(P)}{x_1} \leqq \operatorname*{Max}_{Q \in S_r} \frac{\overset{+}{u}(Q)}{x_1} \leqq \operatorname*{Max}_{Q' \in S_{r'}} \frac{\overset{+}{u}(P')}{x_1'}$$

gilt.

Für den Fall $n = 2$ und einer in der rechten Halbebene $x > 0$ regulären analytischen Funktion $f(z)$ der komplexen Veränderlichen $z = x + iy$ liefert dieser Satz den klassischen Satz von Phragmén und Lindelöf in folgender scharfer Form:

Ist $f(z)$ in $x>0$ regulär analytisch, und gilt für jeden endlichen Punkt ζ der imaginären Achse

$$\lim_{z\to\zeta}|f(z)|\leq 1 \tag{5.2.4}$$

so gilt entweder $|f(z)|\leq 1$ in $x>0$ oder der Ausdruck $\overset{+}{\log}|f(z)|:x$, der dort nicht identisch verschwindet, genügt für Halbkreisgebiete dem Maximumprinzip.

Diese Aussage kann auch so formuliert werden:

Die Funktion

$$\varphi(r)=\operatorname*{Max}_{z\in S_r}\{\overset{+}{\log}|f(z)|:x\} \tag{5.2.6}$$

ist eine nicht abnehmende Funktion von r und konvergiert infolgedessen für $r\to\infty$ gegen eine (endliche oder unendliche) Zahl $a>0$.

§6. Wachstum spezieller Funktionenklassen.

6.1. Erweiterung der Klasse L_0. Genügt eine Funktion $u=u(P)$ in H^n den zwei Bedingungen von 1.1, so kann sie dort nach dem eben bewiesenen Satz von Phragmén und Lindelöf entweder beschränkt sein oder mindestens vom Normaltypus von der Ordnung Eins. Durch eine Lockerung der Bedingungen soll es nun möglich werden, die Spanne dieser Alternativen zu verringern.

Eine in H^n reelle eindeutige Funktion $u=u(P)$ soll dort von der Klasse $\overline{L}_0$ heißen, wenn sie folgenden Bedingungen genügt.

1. Die Menge $E(u>0)$ ist nicht leer.

2. In der Umgebung jedes Punktes $P\in H^n$ ist u zweimal stetig differenzierbar.

3. In jedem Punkt der Menge $E(u\geqq 0)$ genügt u der Differentialungleichung

$$\sum_1^n \frac{\partial^2 u}{\partial x_k^2}\geq \frac{c}{r^2}u, \tag{6.1.1}$$

wobei $c=c(P)$ eine in H^n eindeutige stetige Funktion, ist, die der Ungleichung

$$c(P)+n-1\geq 0 \tag{6.1.2}$$

genügt.

4. Die Funktion u genügt der (Lindelöfschen) Bedingung

$$\overline{\lim_{P\to Q}}\,u(P)\leq 0 \qquad \big(P\in H^n,\ Q\in E(x=0)\big). \tag{6.1.3}$$

Infolge der Bedingung 3 braucht hier u keine subharmonische Funktion zu sein. Es wird jedoch gezeigt, daß sie durch die Einschränkung (6.1.2) Eigenschaften beibehält, die sie sehr nahe an die Funktionen der Klasse L_0 bringen.

6.2. Ungleichungen. Setzt man in (4.1.3) $U=V=u$ $(u\in L_0)$, so kann man ohne weiteres zeigen, daß die Punktmenge $E(u\geqq 0)$ keine Inselgebiete enthalten kann, auf deren Rand Γ durchweg $u=0$ gilt. Ist nämlich G ein

solches im Endlichen liegendes Gebiet, so wird nach (4.1.3), wenn ν die Richtung der äußeren Normale bedeutet,

$$\int_{\Gamma} u \frac{\partial u}{\partial \nu} dO_{n-1} = \int_{G+\Gamma} \{u \operatorname{div} \operatorname{grad} u + (\operatorname{grad} u)^2\} d\tau_n.$$

Führt man rechts Polarkoordinaten ein, so wird die geschweifte Klammer mit Rücksicht auf (6.1.1) $\geq u_r^2 + \frac{1}{r^2}\{V(u,u) + c\,u^2\}$. Verwendet man jetzt das Wirtingersche Lemma (3.2.3) ($k=1$!) und (2.3.5) auf jedem Schnitt E_r von G mit der Halbkugel S_r, wonach

$$\int_{E_r} V(u,u)\, d\omega_n \geq (n-1) \int_{E_r} u^2\, d\omega_n \tag{6.2.1}$$

ist, so wird mit Rücksicht auf (6.1.2)

$$\int_{\Gamma} u \frac{\partial u}{\partial \nu} dO_{n-1} \geq \int_{G} u_t^2\, t^{n-1}\, d\omega_n\, dt. \tag{6.2.2}$$

Ist nun $u=0$ auf Γ, so müssen in G folgende Bedingungen erfüllt sein:

1. $$\int_{E_r} V(u,u)\, d\omega_n = (n-1) \int_{E_r} u^2\, d\omega_n$$

für alle r, die den Punkten von G entsprechen.

2. $$c(P) + n - 1 = 0$$

3. $$u_r = 0 \quad (P \in G)$$

4. $$\sum_1^n \frac{\partial^2 u}{\partial x_k^2} = \frac{c}{r^2} u.$$

Wegen 3. und 2. schreibt sich diese Bedingung in der Form

$$u_{rr} + \frac{n-1}{r} u_r + \frac{1}{r^2} \Delta u = \frac{c}{r^2} u. \tag{6.2.3}$$

Dabei ist Δ, wie nochmals wiederholt werden darf, der zweite Beltramische Operator auf der Einheitskugel S.

Nun folgt aus 1. in Verbindung mit der Stetigkeit von u, daß E_r mit S_r übereinstimmen muß. Wegen 3. kann also u nur von der Form $C \cos\vartheta$ mit einem konstanten C sein. In diesem Falle ist dann (6.2.3) automatisch erfüllt.

Man bilde jetzt mit $v=u$ das Analogon von (4.1.7) und beachte, daß, falls dieser Ausnahmefall, der uns im folgenden nicht weiter interessieren wird, eintritt, von einem r an die rechte Seite der Gleichung

$$r^{n-1} \int_{E_r} u \frac{\partial u}{\partial r} d\omega_n = \int_{G_r} f(P)\, t^{n-1}\, dt\, d\omega_n \tag{6.2.4}$$

mit

$$f(P) = u\left\{\frac{1}{r^{n-1}} (r^{n-1} u_r)_r + \frac{1}{r^2} \Delta u\right\} + \left\{u_r^2 + \frac{1}{r^2} V(u,u)\right\}$$

positiv ist. Es genügt also hier

$$m(r) = \int_{S_r} \overset{+}{u}^2 \, d\omega_n \tag{6.2.5}$$

zu setzen, um nach dem Vorbild von 4.1 die Carlemansche Differentialgleichung zu bilden. Das soll nun in der nächsten Nummer geschehen.

6.3. Die Fundamentalungleichung. Man wiederhole hier nun den Gedankengang von 4.1 unter Berücksichtigung der veränderten Voraussetzungen über u, indem man hier wieder mit dem Gebiet G_r operiert.

Nimmt man r groß genug, damit $m > 0$ ausfällt, so wird

$$\frac{1}{r^{n-1}} (r^{n-1} m')' \geqq \frac{1}{2} \frac{m'^2}{m} + \frac{2}{r^2} \int_{\Gamma_r} c\, u^2 \, d\omega_n + \frac{2}{r^2} \int_{\Gamma_r} V(u, u) \, d\omega_n .$$

Durch Anwendung des Wirtingerschen Lemmas läßt sich diese Ungleichung auch in der Form schreiben

$$\frac{1}{r^{n-1}} (r^{n-1} m')' \geqq \frac{1}{2} \frac{m'^2}{m} + \frac{2(n-1)}{r^2} m + \frac{2}{r^2} \int_{\Gamma_r} c\, u^2 \, d\omega_n .$$

Wie in dem Fall einer subharmonischen Funktion kann m'' unter Umständen die untere Derivierte von m' bedeuten. Das bleibt jedoch auf die Frage der Konvexität ohne Einfluß. Es sei nun

$$q = q(r) = \operatorname*{Min}_{P \in S_r} c(P) . \tag{6.3.1}$$

Dann wird

$$\frac{1}{r^{n-1}} (r^{n-1} m')' \geqq \frac{1}{2} \frac{m'^2}{m} + \frac{2(q+n-1)}{r^2} m . \tag{6.3.2}$$

Diese Ungleichung, die von einem r an gilt, schränkt das Wachstum der Funktionen der Klasse $\overline{L}_0$ ein. Der nächste Paragraph soll der Entwicklung der Konsequenzen von (6.3.2) gewidmet werden.

§ 7. Konvexitätssätze und Phragmén-Lindelöfsches Prinzip für die Klasse L_2. Der Hadamardsche Dreikugelsatz.

7.1. Die Klasse $\overline{L}_2$ und die Konvexitätssätze. Wie beim Fall der subharmonischen definiert man hier die Klasse $\overline{L}_2$ durch Hinzufügung der Grenzfunktionen aller monoton abnehmenden, konvergenten Folgen von Funktionen der Klasse $\overline{L}_0$. Die Klasse, welche aus allen gleichmäßig konvergenten monoton abnehmenden Folgen von Funktionen der Klasse $\overline{L}_0$ besteht, wird entsprechend mit $\overline{L}_1$ bezeichnet. Jedes $u \in \overline{L}_1$ ist dann in H^n stetig.

Infolge der Veränderlichkeit von $q(r)$ können Konvexitätssätze für die Klassen $\overline{L}_0$, $\overline{L}_1$ und $\overline{L}_2$ nur für feste Intervalle $J(r_1 < r < r_2)$ gegeben werden.

Es sei $q(J)$ durch die Gleichung

$$q(J) = q = \inf_{r \in J} q(r) \tag{7.1.1}$$

definiert und $\lambda = \lambda(J)$ durch die Gleichung

$$\lambda(\lambda + n - 1) = q + n - 1 \tag{7.1.2}$$

bestimmt.

Setzt man dann $\xi = r^{2\lambda+n-2}$ und

$$y(\xi) = m(r)\, r^{n-2} \tag{7.1.3}$$

so transformiert sich (6.3.2) wegen

$$\frac{m'}{m} = (2\lambda + n - 2)\,\frac{\dot{y}}{y}\, r^{2\lambda+n-3} - \frac{n-2}{r}$$

und

$$\frac{m''}{m} + \frac{n-1}{r}\frac{m'}{m} = (2\lambda + n - 2)^2\,\frac{\ddot{y}}{y}\, r^{2(2\lambda+n-3)} + 2\lambda(2\lambda + n - 2)\,\frac{\dot{y}}{y}\, r^{2\lambda+n-4}$$

in die Ungleichung

$$\frac{d^2 y}{d\xi^2} \geq \frac{1}{2}\, y \left(\frac{d}{d\xi} \log \frac{y}{\xi}\right)^2. \tag{7.1.4}$$

Es gilt also der Satz:

Ist $r \in J$ und wird λ durch (7.1.2) definiert, so ist dort $m(r)\, r^{n-2}$ eine konvexe Funktion von $r^{n+2\lambda-2}$.

Schreibt man wieder g für $m(r)^{\frac{1}{2}} : r^{\lambda}$, so genügt g der Differentialungleichung

$$g'' + \frac{n + 2\lambda - 1}{r}\, g' \geq 0, \tag{7.1-5}$$

aus der, ähnlich wie vorhin, der Satz folgt:

Die Funktion $m^{\frac{1}{2}} : r^{n+2\lambda-2}$ ist in J eine konvexe Funktion von $1 : r^{n+2\lambda-2}$.

Schreibt man die Konvexitätseigenschaften in der Form der Ungleichung (5.1.1), so kann man durch Grenzübergang die beiden eben bewiesenen Sätze für alle Funktionen der Klasse $\overline{L}_2$ beweisen. Wir gehen darauf nicht näher ein.

Die beiden Sätze sind insofern genau, als sie auch den linearen Fall einschließen. Das zeigen die Beispiele $u r^{\alpha} = \cos\vartheta$ $(\alpha > 0)$.

7.2. Nochmals das Phragmén-Lindelöfsche Prinzip. Ist $u \in L_0$, so definiere man die Folge

$$q_m = \inf_{r \geq m} \{q(r)\} \qquad (m = 1, 2, \ldots) \tag{7.2.1}$$

und setze

$$q_0 = \lim_{m \to \infty} q_m. \tag{7.2.2}$$

Schreibt man dann λ_m für $\lambda(q_m)$ und λ_0 für $\lambda(q)$, so kann man zeigen:

1. Für jedes m existieren stets die Grenzwerte

$$\lim_{r \to \infty} m(r)^{\frac{1}{2}} : r^{q_m} = A_m \qquad (m = 1, 2, \ldots). \tag{7.2.3}$$

2. Für jedes m ist $A_m > 0$.

In der Tat existieren wegen der Konvexität von $m(r)\, r^{n-2}$, in Verbindung mit der Tatsache $m' > 0$, die nichtnegativen Grenzwerte

$$\lim_{r\to\infty} \frac{m(r)\, r^{n-2} - m(r_1)\, r_1^{n-2}}{r^{2\lambda_m+n-2} - r_1^{2\lambda_m+n-2}} = C_m \qquad (m = 1, 2, \ldots) \tag{7.2.4}$$

und infolgedessen auch (durch Festhalten von r_1) die Grenzwerte

$$A_m = \lim_{r\to\infty} m(r) : r^{2\lambda_m} = C_m. \tag{7.2.5}$$

Das beweist die Behauptung.

Die Folge (λ_m) ist ähnlich wie die Folge (q_m) nicht abnehmend und infolgedessen kann man im allgemeinen keine Aussage über das Verhalten des Ausdrucks $m(r) : r^{2\lambda_0}$, falls $r \to \infty$ konvergiert, machen. Die einzige allgemeine Aussage, die man hier machen kann, ist die, daß nämlich stets

$$\lim_{r\to\infty} \frac{\log m(r)^{\frac{1}{2}}}{\log r} \geqq \lambda_0 \tag{7.2.6}$$

gilt, also auch

$$\lim_{r\to\infty} \frac{\log M(r)}{\log r} \geqq \lambda_0. \tag{7.2.7}$$

Hierbei bedeutet, wie üblich, $M(r)$ das Maximum von $\overset{+}{u}(P)$ auf S_r. Die Zahl $\lambda_0 \left(\lambda_0 = \lim_{m\to\infty} \lambda(q_m)\right)$ kann hier alle nichtnegativen Werte annehmen und insofern bedeutet die Aussage (7.2.7) eine Vertiefung des Phragmén-Lindelöfschen Prinzips, das lediglich die Werte $\lambda_0 \geqq 1$ zuläßt.

Wie die Funktionen

$$u(P) = \left(r^{\alpha} \prod_1^e \log_k^{\alpha_k} r\right) \cos\vartheta \qquad (\alpha_k \gtrless 0),$$

zeigen, für die

$$c(P) = \alpha(\alpha + n - 2) - (n-1) + o(1)$$

gilt, und insofern in (7.2.7) das Gleichheitszeichen eintritt, ist kaum zu erwarten, wie im Falle der harmonischen Funktionen, eine Aussage über den Typus von u zu erhalten.

Im Falle $\lambda_0 > 0$ ist das Resultat (7.2.7) insofern bemerkenswert, als es zeigt, daß man durch Hinzunahme der Klasse der subharmonischen Funktionen mit $-(n-1) \leqq c(P) \leqq 0$ die Spanne im Phragmén-Lindelöfschen Prinzip (beschränkt oder mindestens von der Ordnung Eins) beliebig verringern kann.

7.3. Der Hadamardsche Dreikugelsatz in R^n. Die Klasse S_2 der subharmonischen Funktionen von R^n besteht bekanntlich aus allen zweimal stetig differenzierbaren subharmonischen Funktionen (Klasse S_0) von R^n und

allen nicht zunehmenden, konvergenten Folgen von S_0. Früheren Bezeichnungen entsprechend bezeichnen wir die Klasse aller stetigen subharmonischen Funktionen, die Limes von monoton abnehmenden, gleichmäßig konvergenten subharmonischen Funktionen der Klasse S_0 sind, mit S_1.

Nachfolgender Satz ist nicht tief und soll hier nur der Vollständigkeit wegen gebracht werden. Ist $u \in S_0$ und bildet man die Mittelwerte

$$\varrho_k(r) = \int_{S_r} u^+(P)^{2k}\, d\omega_n, \tag{7.3.1}$$

wobei S_r hier die volle Kugel $x_1^2 + \cdots + x_n^2 = r^2$ bedeutet, so kann man leicht die Differentialungleichung

$$\frac{1}{r^{n-1}} (r^{n-1} \varrho_k')' \geq \frac{2k-1}{2k} \frac{\varrho_k'^2}{\varrho_k}, \tag{7.3.2}$$

d.h.

$$\frac{\varrho_k''}{\varrho_k} + \frac{n-1}{r} \frac{\varrho_k'}{\varrho_k} - \frac{2k-1}{2k} \left(\frac{\varrho_k'}{\varrho_k}\right)^2 \geq 0 \tag{7.3.3}$$

gewinnen. Setzt man hier $\mu_k = \varrho_k^{1:2k}$, so wird nach leichten Umrechnungen

$$(r^{n-1} \mu_k')' \geq 0. \tag{7.3.4}$$

Ist nun $n = 2$, so drückt diese Ungleichung die Konvexität von μ_k als Funktion von $\log r$, und zwar (nach geeignetem Grenzübergang) für jedes $u \in S_0$. Drückt man wieder die Konvexität von μ_k in der Form (5.1.1) aus, und geht man mit $k \to \infty$, so erhält man die Behauptung für $M(r)$ für die Klassen S_0 und S_1. Die Klasse S_2 erledigt sich durch die schon vorhin gegebene Erweiterung der Definition von $M(r)$. Ist nun $n > 2$ und setzt man $y = \mu(r)\, r^{n-2}$, $\xi = r^{n-2}$, so erhält man aus (7.3.4)

$$\frac{y''}{y} - \frac{n-3}{r} \frac{y'}{y} \geq 0,$$

d.h. $\frac{d^2 y}{d\xi^2} \geq 0$. Das Übrige, die Konvergenz $k \to \infty$ und der Übergang zu den Klassen S_1, S_2 erledigt sich wie vorhin. Zusammenfassend können wir den Dreikugelsatz von Hadamard in folgender Form beweisen:

Es sei $n > 2$ und $u \in S_2$. Definiert man für jedes r die Größe $M(r)$ als die untere Grenze der Zahlen M, für die $E(u > M,\ P \in S_r)$ das Maß Null hat, so ist $M(r)\, r^{n-2}$ eine konvexe Funktion von r^{n-2}.

7.4. Nichtlineare Fälle. Um zu zeigen, daß die Klasse S_2 eine ziemlich abgeschlossene Klasse darstellt, will ich, ohne damit Vollständigkeit anzustreben, zeigen, daß unter bestimmten einfachen Wachstumsvoraussetzungen die Differentialgleichung

$$\sum_1^n \frac{\partial^2 u}{\partial x_k^2} = g\big(P, u(P)\big), \tag{7.4.1}$$

sobald g in u nicht linear ist, keine durchweg in R_n regulären Integrale besitzen kann. Der Fall einer Ungleichung erledigt sich dann analog.

Der Einfachheit halber setze ich $n=2$ voraus und zeige:

Gilt bei vorgegebenem $\varepsilon>0$ von einem $r=r(\varepsilon)$ an

$$g\left(P,\overset{+}{u}(P)\right)\geq\frac{C}{r^2}\overset{+}{u}^{1+4\varepsilon}\qquad\left(C=C(\varepsilon)\right), \tag{7.4.2}$$

so kann (7.4.1) keine im Endlichen regulären Integrale haben.

Zunächst kann man ohne weiteres annehmen, daß es mindestens einen Punkt P_0 gibt, in dem $u(P_0)>0$ ist. Dann zeigt man, daß $E(u>0)$ keine im Endlichen liegenden Gebiete enthalten kann, auf deren Rand $u=0$ ist.

Man setze nun zunächst $u\in S_0$ voraus und verwende wieder den Ausdruck

$$m(r)=\frac{1}{2\pi}\int_0^{2\pi}\overset{+}{u}^2(P)\,d\varphi\qquad(P\in|z|=r). \tag{7.4.3}$$

Setzt man hier $r\geq r(\varepsilon)$ voraus, so wird wegen

$$\frac{1}{2\pi}\int_0^{2\pi}\overset{+}{u}^{2+4\varepsilon}\,d\varphi\geq m^{1+2\varepsilon},$$

$$\left(\frac{m''}{m'}+\frac{1}{r}\right)^2\geq 2\left(\frac{m''}{m'}+\frac{m'}{r\,m}\right)-\left(\frac{m'}{m}\right)^2\geq\frac{C}{r^2}m^{2\varepsilon}. \tag{7.4.4}$$

Daß hier m von einem r an positiv ist, und daß links der eingeklammerte Ausdruck nicht negativ ist, sieht man aus den Entwicklungen von 4.1.

Aus (7.4.4) folgt nun für $r\geq r(\varepsilon)$

$$\frac{m''}{m'}+\frac{1}{r}\geq\frac{C_1}{r}m^{\varepsilon}\qquad(C_1=C_1(\varepsilon)>0)$$

oder wenn $r=e^{\xi}$ gesetzt wird,

$$\frac{d}{d\xi}\log\frac{dm}{d\xi}\geq C_1 m^{\varepsilon}.$$

Das gibt die Ungleichung

$$\frac{d^2m}{d\xi^2}\geq C_2\frac{dm^{1+\varepsilon}}{d\xi},$$

aus der dann durch Integration die Ungleichung

$$\int^{r}\frac{dm}{m^{1+\varepsilon}}\geq C_2\int^{r}\frac{dr}{r}+C_3\qquad(C_2>0) \tag{7.4.5}$$

mit einem konstanten C_2, C_3 folgt. Das beweist die Behauptung. Der allgemeine Fall $u\in S_2$ kann durch Grenzübergang erledigt werden.

Der Schluß funktioniert auch dann, wenn $u \in M_0$ ist und auf einer Teilmenge E_1 von $E(u > 0)$, die als nicht leer vorausgesetzt wird, der Ungleichung

$$\sum_1^n \frac{\partial^2 u}{\partial x_k^2} \geq \frac{C}{r^2} u \qquad (r \geq r(\varepsilon),\ C = C(\varepsilon) < 0) \tag{7.4.6}$$

genügt.

Diese Teilmenge E_1 soll folgende Eigenschaften besitzen:

1. Ist $P \in E_1$ und liegt P auf der Kugel K_r, $x_1^2 + \cdots + x_n^2 = r^2$, so liegt der Durchschnitt $K_r \cap E$ in E_1.

2. Es sei $\varphi(x)$ die charakteristische Funktion der Menge E_1^* der r-Achse, die alle Radien der K_r enthält, die Punkte von E_1 enthalten.

Dann ist $\int_{E_1^*} \varphi(x) \frac{dx}{x}$ divergent.

Wiederholt man dann die eben durchgeführten Rechnungen, so wird

$$r\left(\frac{m''}{m'} + \frac{1}{r}\right) \geq C_1 \varphi(r)\, m^\varepsilon, \tag{7.4.7}$$

d.h.

$$\frac{d}{d\xi} \log \frac{dm}{d\xi} \geq C_1 \varphi(r)\, m^\varepsilon. \tag{7.4.8}$$

Daraus folgt für jedes $\xi \geq \xi(\varepsilon)$

$$\frac{d^2 m}{d\xi^2} \geq C_2 \varphi \frac{dm^{1+\varepsilon}}{d\xi},$$

und infolgedessen für jedes ξ mit $e^\xi \in E_1^*$

$$\frac{dm}{d\xi} \geq C_1 \varphi(\xi)\, m^{1+\varepsilon}. \tag{7.4.9}$$

Denn sowohl m als auch $dm/d\xi$ sind monoton wachsend und letztere bei geeigneter Wahl von $\xi(\varepsilon)$ positiv.

Aus (6.4.6) erhält man jetzt für jedes $r \in E_1^*$

$$\varphi(r) \frac{dr}{r} \leq C_2 \frac{dm}{m^{1+\varepsilon}}, \tag{7.4.10}$$

woraus dann durch Integration die Behauptung folgt.

Ich will die Zahl der Beispiele der Anwendungsmöglichkeit der Carlemanschen Methode nicht weiter erhöhen. Man könnte natürlich auch manche Probleme aus dem Lindelöfschen Kreis verschärfen, indem man z.B. annimmt, daß die Lindelöfsche Bedingung $u \leq 0$ nicht überall erfüllt ist, sondern auf Gebieten von $x_1 = 0$, in welchen eine unendliche Folge von monoton wachsenden Ringgebieten Platz findet, die sich ins Unendliche erstrecken. Auch könnte man einen Vorstoß zu denjenigen Lindelöfschen Problemen unternehmen, welche durch die allgemeine elliptische partielle Differentialgleichung eröffnet werden. Einen guten Anfang in dieser Richtung hat Huber gemacht (Proc. Amer. Math. Soc. **4**, 852–857, 1953).

Andere Probleme, die sich mit der hier entwickelten Carlemanschen Methode angreifen lassen, bewegen sich in der Richtung des Denjoy-Carleman-Ahlforsschen Satzes. In dieser Richtung liegt, soviel ich weiß, nur ein Ergebnis vor, das ich in einer kurzen Comptes Rendus Note aus dem Jahre 1954 angegeben habe. Erst wenn man diese Resultate hinzunimmt, erhält man einen Einblick über die Menge der Probleme, in die man durch die hier entwickelte Methode vordringen kann. Ich glaube jedoch, daß der Leser durch die hier zur Darstellung gekommenen Auswahl der Anwendungen der verallgemeinerten Carlemanschen Methode einen überzeugenden Eindruck von der Stärke der ursprünglichen Carlemanschen Idee erhalten hat.

Berlin-Dahlem, Mathematisches Institut der Freien Universität.

(Eingegangen am 11. Januar 1955.)

Soeben erschien:

Physik der Sternatmosphären

mit besonderer Berücksichtigung der Sonne

Von Dr. **Albrecht Unsöld**, Professor für theoretische Physik an der Universität Kiel. Zweite Auflage. Mit 257 Abbildungen. IX, 866 Seiten Gr.-8°. 1955.

Ganzleinen DM 168.—

Inhaltsübersicht: **I. Die Sternatmosphäre im thermischen Gleichgewicht.** Strahlungstheorie. — Anwendung des PLANCKschen Gesetzes auf die Strahlung der Sonne. Strahlungsmessungen. — Die Strahlung der Sterne. — Thermische Ionisation und Anregung. — **II. Kontinuierliches Spektrum und Aufbau einer Sternatmosphäre.** Strahlungsaustausch. — Strahlungsgleichgewicht. — Kontinuierliche Absorption und Streuung. — Der Aufbau der Sternatmosphären. — **III. Messung der Intensitätsverteilung in den FRAUNHOFER-Linien.** Messung der Profile und Äquivalentbreiten von FRAUNHOFER-Linien. — **IV. Physikalische Grundlagen der Theorie der FRAUNHOFER-Linien.** Dämpfung (insbesondere Strahlungsdämpfung) und Dopplereffekt. — Druckverbreiterung und Druckverschiebung von Spektrallinien. — Quantenmechanische Berechnung von Linienstärken, Übergangswahrscheinlichkeiten und Oszillatorenstärken. — Experimentelle Bestimmung von Oszillatorenstärken. Laboratoriums-Untersuchungen über Strahlungsdämpfung sowie Druckverbreiterung und -verschiebung von Spektrallinien. — Literatur über astrophysikalisch wichtige Linienstärken, Übergangswahrscheinlichkeiten und Oszillatorenstärken einzelner Elemente. — **V. Die Entstehung der FRAUNHOFER-Linien.** Strahlungsaustausch und FRAUNHOFER-Linien. — Theorie der Wachstumskurven. — Quantitative Deutung der FRAUNHOFER-Linien in den Spektren der Sonne und „normaler" Sterne. — Beeinflussung der Linienprofile durch Rotation und Expansion der Sterne. — Klassifikation der Sternspektren. — **VI. Physik der Sonne.** Die Struktur der äußeren Schichten der Sonne. — Der Sonnenrand. Chromosphäre, Korona und Protuberanzen. — **VII. Radiofrequenzstrahlung und kosmische Ultrastrahlung.** Radioastronomie. — Kosmische Ultrastrahlung und Radiofrequenzstrahlung. — **Anhang:** Naturkonstanten und Zahlenwerte. — Klassifikation der Linienspektren. — Über die Integralexponentialfunktionen. — Literaturverzeichnis. — Namen- und Sachverzeichnis.

Die glänzende Entwicklung der Astrophysik in den letzten 20 Jahren machte eine weitgehende Neubearbeitung der „Physik der Sternatmosphären" erforderlich. In der Theorie des Strahlungsgleichgewichtes wie auch des Energietransportes durch Konvektion sind wichtige neue Methoden entwickelt worden. Die Quantentheorie der kontinuierlichen Absorption und Streuung sowie eine eingehende Behandlung von Dopplereffekt, Strahlungsdämpfung, Stoßdämpfung, Druckverbreiterung, Berechnung und Messung von Übergangswahrscheinlichkeiten etc. dienen zur Vorbereitung der Theorie der Spektren von Sonne und Sternen. Die ganzen Probleme der Linienprofile, Wachstumskurven, Atmosphärenmodelle, der chemischen Zusammensetzung der Sterne, Klassifikation der Sternspektren etc. werden ihrem heutigen Stand entsprechend vorgeführt.

Weitere Kapitel sind der Physik der Sonne (Granulation, Chromosphäre und Korona) und den besonders aktuellen Problemen der Radioastronomie und den astrophysikalischen Aspekten der kosmischen Ultrastrahlung gewidmet. Den Schluß bildet ein Literaturverzeichnis von über 50 Seiten.

Das Buch dürfte nicht nur für Astronomen und Astrophysiker nützlich sein; es enthält auch vielerlei Hilfsmittel und Anregungen für die Gebiete der Spektroskopie und Gasentladungsphysik, der kosmischen Ultrastrahlung und Radioastronomie sowie der Ionosphärenforschung und Meteorologie.

SPRINGER-VERLAG / BERLIN · GÖTTINGEN · HEIDELBERG

Darstellungen von Gruppen

mit Berücksichtigung der Bedürfnisse der modernen Physik

Von Dr. **Hermann Boerner**, Professor der Mathematik an der Justus-Liebig-Hochschule, Gießen. Mit 15 Abbildungen. XI, 287 Seiten Gr.-8°. 1955. (Die Grundlehren der mathematischen Wissenschaften, Band LXXIV.)

DM 33.—; Ganzleinen DM 36.60

Inhaltsübersicht: **I. Matrizen.** Vektoren. Lineare Abbildungen. Matrizen. Begriff einer Algebra. Quadratische und hermitesche Formen, orthogonale und unitäre Matrizen. Eigenwerte und Transformation auf Diagonalgestalt. Zwei weitere Verknüpfungen für Matrizen; das KRONECKER-Produkt. Äquivalenz und Reduzibilität von Matrixsystemen. Das Lemma von SCHUR. Vertauschbarkeit von Matrixsystemen. Beispiele irreduzibler Systeme. Eine Anwendung des SCHURschen Lemmas. — **Gruppen.** Elementare Gruppentheorie. Die symmetrische und die alternierende Gruppe. Kontinuierliche Gruppen. Die Matrix-Exponentialfunktion. Der Infinitesimalring einer linearen Gruppe. Integration in LIEschen Gruppen. — **Allgemeine Darstellungstheorie.** Begriff der Darstellung. Die vollständige Reduzibilität der Darstellungen endlicher Gruppen. Eindeutigkeit der Zerlegung. Der Gruppenring und die reguläre Darstellung. Struktur des Gruppenrings. Vorbereitende Sätze. Die Struktur des Gruppenrings und das System der Klassen irreduzibler Darstellungen. Zur Darstellungstheorie der halbeinfachen Algebren. Normale Darstellungen. Die Charaktere. Charaktere und Gruppenring. Darstellungen und Charaktere eines direkten Produkts. Zusammenhang der Charaktere mit denen einer Untergruppe. Weitere Formeln für die Charaktere; ihre Berechnung auf algebraischem Wege. Die infinitesimalen Transformationen der Darstellungen kontinuierlicher Gruppen. Die adjungierte Darstellung. Die Charaktere der kontinuierlichen Gruppen. Gruppen mit Normalteiler vom Index 2. — **Die Darstellungen der symmetrischen Gruppen.** Die Tableaux. Hilfssätze über die Tableaux. Die irreduziblen Darstellungen. Die Standard-Tableaux. Volle Reduktion des Gruppenrings. Berechnung der Matrizen einer irreduziblen Darstellung. Beweis der Sätze 4.2 und 4.3. — **Die Darstellungen der vollen linearen, unimodularen und unitären Gruppen.** Vorbemerkungen. Das KRONECKER-Quadrat und die symmetrischen und schiefsymmetrischen Tensoren zweiter Stufe. Der Raum der Tensoren ν-ter Stufe und die Darstellungen der Gruppe G_n vom Polynomgrad ν. Die Symmetrieklassen im Tensorraum. Die Tableaux und die ganzrationalen Darstellungen der vollen linearen Gruppe. Der Verzweigungssatz. Ganzrationale Darstellungen der reellen linearen, unimodularen und unitären Gruppen. Rationale und semirationale Darstellungen. Die unzerfällbaren Darstellungen der additiven Gruppe der reellen Zahlen. Die stetigen Darstellungen der vollen und reellen linearen, der unimodularen und unitären Gruppen. — **Charaktere der linearen und der Permutationsgruppen. Die alternierende Gruppe.** Die Charakteristiken und die Darstellungsgrade der ganzrationalen Darstellungen der vollen linearen Gruppe. Zusammenhang zwischen den Charakteren der symmetrischen Gruppe und den Charakteristiken der vollen linearen Gruppe. Zur Berechnung der Charaktere der symmetrischen Gruppe. Übersicht über die Darstellungen der alternierenden Gruppe. Noch eine Formel zur Berechnung der Charaktere von $\mathfrak{S}_p$. Analyse von KRONECKER-Produkten bei der symmetrischen und bei der vollen linearen Gruppe. Die Charaktere der alternierenden Gruppe. — **Charaktere und eindeutige Darstellungen der Drehgruppe.** Zusammenhangsverhältnisse der Drehgruppe. Das Toroid $\mathfrak{T}_p$. Das STIEFELsche Diagramm. Die Gruppe Ψ. Die Fundamentalbereiche der Gruppe Ψ. Die Eigenwerte der Darstellungen. Die Eigenwerte der adjungierten Darstellung. Das Integral über eine Klassenfunktion. Invariante und alternierende Polynome und Elementarsummen. Das System der einfachen Charaktere. Der Darstellungsgrad. Der Verzweigungssatz. Anwendung auf die niederen Dimensionszahlen. Die Fundamentaldarstellungen. Die volle orthogonale Gruppe. — **Spindarstellungen, Infinitesimalring, gewöhnliche Drehgruppe.** Der Infinitesimalring der Drehgruppe. CLIFFORDs Algebra und ihr Zusammenhang mit den infinitesimalen Drehungen. Darstellungstheorie der CLIFFORDschen Algebra. Die Spindarstellungen des Infinitesimalrings der Drehgruppe. Die Spindarstellungen der Drehgruppe. Die gewöhnliche Drehgruppe $\mathfrak{d}_3$. Die Formel von CLEBSCH-GORDAN. Struktur des Infinitesimalrings und Gewichte der Darstellungen. Weitere KRONECKER-Produkte. Algebra von KEMMER und DE BROGLIE. — **Die LORENTZ-Gruppe.** Die vier Stücke der LORENTZ-Gruppe. Die Fundamentaldarstellungen der LORENTZ-Gruppe $\mathfrak{L}_{n,i}$. Die gewöhnliche eigentliche LORENTZ-Gruppe $l_{4,1}$ und ihr Zusammenhang mit der unimodularen Gruppe g_2. Die Darstellungen der vollen LORENTZ-Gruppe $\mathfrak{L}_{4,1}$. — Literaturverzeichnis. — Namen- und Sachverzeichnis.

SPRINGER-VERLAG / BERLIN · GÖTTINGEN · HEIDELBERG

Diesem Heft liegen zwei Prospekte des Springer-Verlages, Berlin - Göttingen - Heidelberg, bei.

Springer-Verlag, Berlin · Göttingen · Heidelberg. — Druck der Universitätsdruckerei H. Stürtz AG., Würzburg.
Printed in Germany.